



**HAL**  
open science

# CRITIQUE OF IMPURE REASON

Steven James Bartlett

► **To cite this version:**

Steven James Bartlett. CRITIQUE OF IMPURE REASON: Horizons of Possibility and Meaning. Studies in Theory and Behavior, 2020. hal-02923307v1

**HAL Id: hal-02923307**

**<https://hal.science/hal-02923307v1>**

Submitted on 27 Aug 2020 (v1), last revised 7 Sep 2021 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

**A Revolutionary Paradigm Shift in  
Philosophical Thought**

**CRITIQUE OF  
IMPURE REASON  
HORIZONS OF POSSIBILITY  
AND MEANING**

**STEVEN JAMES BARTLETT**

*With a Foreword by  
C. F. von Weizsäcker*

**STUDIES IN THEORY AND BEHAVIOR**

[Occasional pages in this eBook are intentionally left blank]



# CRITIQUE OF IMPURE REASON

HORIZONS OF POSSIBILITY  
AND MEANING

Due to the global pandemic in progress at the time of this book's completion, the already long delays of conventional book publishing have become even longer, while the author's awareness of the finitude of a senior researcher's life has increased. As a result, in cooperation with the publisher, he has decided to make this work available as an open access publication, preferring to forego potential book royalties and, at least for some readers, the possible cachet supplied by a large brand-name publisher.

At a time when the needs and interests of readers are rapidly shifting away from the traditional paper printing of books and are much more efficiently and effectively served by electronic publications, it is hoped that readers will welcome the free, open access publication of a book of this size for which a commercial press would need to set a high price, limiting its distribution and readership.

At a later date the publisher may make a modestly priced printed edition available at cost for readers who continue to prefer traditionally printed books.

# CRITIQUE OF IMPURE REASON

## HORIZONS OF POSSIBILITY AND MEANING

STEVEN JAMES BARTLETT

*With a Foreword by*

CARL FRIEDRICH VON WEIZSÄCKER



**STUDIES IN THEORY AND BEHAVIOR**

AN IMPRINT OF AUTOGRAPH EDITIONS

Salem, Oregon, U.S.A.

## AUTOGRAPH EDITIONS

Established 1975

© 2020 by Steven James Bartlett

The author and the publisher have made this book available as an open access publication, freely available to readers under the terms of the **Creative Commons Attribution-NonCommercial-No-Derivs** license, which allows anyone to distribute this work without changes to its content, provided that both the author and the original URL from which this work was obtained are mentioned, that the contents of this work are not used for commercial purposes or profit, and that this work will not be used without the copyright holder's written permission in derivative works (i.e., you may not alter, transform, translate, or build upon this work without such permission). The full legal statement of this license may be found at

<http://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>



*This book has been made available as an open access publication through the courtesy of the publisher. Autograph Editions is a non-profit press committed to the freedom of non-commercial access to original and important works. The author receives no royalties from this publication and the publisher is not compensated for its work.*

*Studies in Theory and Behavior*

An Imprint of

**AUTOGRAPH EDITIONS**

Salem, Oregon, U.S.A.



DEDICATED TO

**Karen**

My love, wife, best friend, and life companion

and

IN MEMORY OF

**Frederic Brenton Fitch**

who recognized the need for philosophical comprehension on  
the level of maximum theoretical generality,

and

**Henry W. Johnstone, Jr.**

who understood that all philosophical positions  
are inherently framework-relative



*Philosophy, from the earliest times, has made greater claims,  
and achieved fewer results, than any other  
branch of learning.*

– Bertrand Russell (1972/1914, p. 13)



We tend either not to recognize or not to accept that we all-too-often trespass beyond the boundaries of the frameworks that make knowledge possible and the world meaningful.

This is a book about the boundaries of frameworks and about the unrecognized conceptual confusions in which we become entangled by trespassing beyond the limits of the possible and meaningful.

In Kant's *Critique of Pure Reason* we find an analysis of the preconditions of experience and of knowledge.

In contrast, but yet in parallel, in the present study our interest is rather in the ways—unfortunately very widespread and often unselfconsciously habitual—in which many of the concepts that we formulate and the claims that we make using them *conflict* with the very preconditions of meaning and of knowledge.

The objective of this study is, in short, a “*critique of impure reason*.” Its purpose is : first, to enable us to recognize the boundaries of what is referentially forbidden—the limits beyond which reference becomes meaningless—and second, to avoid falling victims to a certain broad class of conceptual confusions that lie at the heart of many major philosophical problems. In the process we shall delimit the domain of *possible meaning*.



## CONTENTS AT A GLANCE

*For chapter contents, see Detailed Table of Contents  
beginning on page xiv.*

Preface	xxix
Foreword by Carl Friedrich von Weizsäcker	xxxiii
Acknowledgments	xxxv
Avant-propos: A philosopher's rallying call	1
Introduction	3
A note to the reader	11
A note on conventions	13

### PART I

#### WHY PHILOSOPHY HAS MADE NO PROGRESS AND HOW IT CAN

		17
1	Philosophical-psychological prelude	19
2	Putting belief in its place: Its psychology and a needed polemic	34
3	Turning away from the linguistic turn: From theory of reference to metalogic of reference	46
4	The stepladder to maximum theoretical generality	67

### PART II

#### THE METALOGIC OF REFERENCE

##### A New Approach to Deductive, Transcendental Philosophy

		85
5	Reference, identity, and identification	87
6	Self-referential argumentation and the metalogic of reference	103
7	Possibility theory	131
8	Presupposition logic, reference, and identification	156

## CONTENTS AT A GLANCE

9	Transcendental argumentation and the metalogic of reference	194
10	Framework relativity	205
11	The metalogic of meaning	225
12	The problem of putative meaning and the logic of meaninglessness	256
13	Projection	280
14	Horizons	302
15	De-projection	323
16	Self-validation	341
17	Rationality: Rules of admissibility	353

### **PART III**

#### **PHILOSOPHICAL APPLICATIONS OF THE METALOGIC OF REFERENCE**

##### **Major Problems and Questions of Philosophy and the Philosophy of Science** **363**

18	Ontology and the metalogic of reference	369
19	Discovery or invention in general problem-solving, mathematics, and physics	376
20	The conceptually unreachable: “ <i>The far side</i> ”	395
21	The projections of the external world, things-in-themselves, other minds, realism, and idealism	421
22	The projections of time, space, and space-time	442
23	The projections of causality, determinism, and free will	459
24	Projections of the self and of solipsism	489
25	Non-relational, agentless reference and referential fields	511
26	Relativity physics as seen through the lens of the metalogic of reference	525
27	Quantum theory as seen through the lens of the metalogic of reference	567
28	Epistemological lessons learned from and applicable to relativity physics and quantum theory	612

CONTENTS AT A GLANCE

	<b>PART IV</b>	
	<b>HORIZONS</b>	<b>629</b>
29	Beyond belief	631
30	<i>Critique of Impure Reason</i> : Its results in retrospect	657
	<b>SUPPLEMENT</b>	
	<b>The Formal Structure of the Metalogic of Reference</b>	<b>675</b>
	<b>APPENDIX I</b>	
	<b>The Concept of Horizon in the Work of Other Philosophers</b>	<b>727</b>
	<b>APPENDIX II</b>	
	<b>Epistemological Intelligence</b>	<b>732</b>
	<i>References</i>	765
	<i>Index</i>	805
	<i>About the author</i>	845

## DETAILED TABLE OF CONTENTS

Preface	xxix
Foreword by Carl Friedrich von Weizsäcker	xxxiii
Acknowledgments	xxxv
Avant-propos: A philosopher's rallying call	1
Introduction	3
A note to the reader	11
A note on conventions	13

### PART I

#### WHY PHILOSOPHY HAS MADE NO PROGRESS AND HOW IT CAN 17

<b>1</b>	<b>Philosophical-psychological prelude</b>	<b>19</b>
1.1	Bifurcations of the human mind: The rational bridge problem	25
1.2	What it takes to change a philosopher's mind	28
<b>2</b>	<b>Putting belief in its place: Its psychology and a needed polemic</b>	<b>34</b>
2.1	Willful blindness	38
2.2	Giving belief its due—i.e., a bad name	39
2.3	The rationale for such a polemic	44
<b>3</b>	<b>Turning away from the linguistic turn: From theory of reference to metalogic of reference</b>	<b>46</b>
3.1	Through the looking glass: The superhighway of philosophy	46
3.2	An exit not taken	48
3.3	The overlooked variety of forms of reference	59
3.4	From theory of reference to its metalogic	64



## DETAILED TABLE OF CONTENTS

<b>4</b>	<b>The stepladder to maximum theoretical generality</b>	<b>67</b>
4.1	Reference as the primary focus of this work	67
4.2	Universality	70
4.3	Invariance	71
4.4	Philosophical neutrality	72
4.5	The level of maximum theoretical generality	73
4.6	Reflexivity on the level of maximum theoretical generality	76
4.7	Expanding the scope of a maximally general theory of theories in order to study preconditions of possibility	77
4.8	The concept of metalogic	78
4.9	Meaning	79
4.10	The self-enclosure of reflexive, maximally general theory	80

## PART II

### THE METALOGIC OF REFERENCE

	<b>A New Approach to Deductive, Transcendental Philosophy</b>	<b>85</b>
--	---	-----------

<b>5</b>	<b>Reference, identity, and identification</b>	<b>87</b>
5.1	What is reference?	87
5.2	Pragmatical reference	88
5.3	The separation of reference from the object of reference	91
5.4	Removing agency, intention, action, and volition—at least provisionally—from theory of reference	92
5.5	Levels of reference and iterative reference	93
5.6	A non-linguistic, non-relational conception of reference	94
5.7	Identification, reference, and coordinate systems	96
5.8	The metalogical study of reference: Preliminary comments	101

## DETAILED TABLE OF CONTENTS

<b>6</b>	<b>Self-referential argumentation and the metalogic of reference</b>	<b>103</b>
6.1	Self-referential argumentation in philosophy	103
6.2	Pragmatical, or performative, self-referential arguments	108
6.3	The critical use of pragmatical self-reference	109
6.4	The constructive use of pragmatical self-reference	110
6.5	Passmore's absolute self-refutation	112
6.6	J. L. Mackie's operational self-refutation	115
6.7	Isaye's rétorsion	119
6.8	Metalogical self-referential arguments	126
6.9	The constructive use of metalogical self-reference	129
 <b>7</b>	 <b>Possibility theory</b>	 <b>131</b>
7.1	The basic vocabulary of the metalogic of reference	131
7.2	The concept of possibility	132
7.3	The spectrum of possibility	132
1.	Psychologically based possibility	132
2.	Temporally based possibility	133
3.	Physical or nomological possibility	134
4.	Epistemic possibility	134
5.	Formal possibility	134
6.	Possibility as complementarity	135
7.	Parametric possibility	136
8.	Many worlds possibility	137
9.	Framework-relative possibility	138
10.	Metalogical possibility: The preconditions of identification	139
7.4	A general theory of possibility	141
7.5	Necessity	142
7.6	Excursus: Modal logic and the present study	144
1.	Obstacles to the philosophical use of formal systems of modal logic	145
2.	Choosing a philosophically appropriate system of modal logic	146
3.	The choice made by C. I. Lewis	147
4.	The choice made by Łukasiewicz	149

## DETAILED TABLE OF CONTENTS

5.	Modal logic appropriate for the metalogic of reference	151
7.7	Summary	154
<b>8</b>	<b>Presupposition logic, reference, and identification</b>	<b>156</b>
8.1	The priority of presuppositions	157
1.	Presuppositions as preliminary assumptions	157
2.	Conditional presuppositions	158
3.	Presuppositions as logical premises	158
4.	Presuppositions as logically antecedent suppositions	158
5.	Presuppositions as the ingredients of definitions	158
6.	Presuppositions as suppositions of language use	159
7.	Presuppositions of pragmatical activity	160
8.	Presuppositions of missing premises	160
9.	Presuppositions of existence	161
10.	Linguistic presuppositions of reference	162
11.	Presuppositions of concepts	163
12.	Structural and systemic presuppositions	163
13.	Presuppositions of identification	163
8.2	Truth-functional presuppositions	165
1.	Necessary and sufficient conditions of truth: Logical implication	166
2.	Entailment	167
3.	Referential presuppositions of truth and falsity	168
8.3	Structural and systemic presuppositions	172
1.	Structural presuppositions	173
2.	Systemic presuppositions	174
8.4	The epistemological loop	176
8.5	Conceptualizing the epistemological loop	182
8.6	Presuppositions of identification, continued	186
Appendix to Chapter 8		
	Rule-based games and Passmore's and Collingwood's presuppositions	191

## DETAILED TABLE OF CONTENTS

<b>9</b>	<b>Transcendental argumentation and the metalogic of reference</b>	<b>194</b>
9.1	What is transcendental argumentation?	194
9.2	Transcendental argumentation as structural/systemic	198
9.3	Transcendental argumentation: Possibility, necessity, and identifiability	200
9.4	Transcendental argumentation and meaning	201
<b>10</b>	<b>Framework relativity</b>	<b>205</b>
10.1	Frames of reference, reference frames, and frameworks	208
10.2	Framework-relative field theory	211
10.3	Framework relativity and ontology	215
10.4	Framework self-enclosure and translation to other frameworks	216
10.5	Framework relativity and perspectives	218
10.6	Framework relativity and conceptual constructs	224
<b>11</b>	<b>The metalogic of meaning</b>	<b>225</b>
11.1	Meaning and theories of meaning	226
11.2	Referential consistency as a criterion of meaning	231
11.3	Referential consistency as an intrinsically determined criterion of meaning	235
11.4	Formalized description	236
11.5	Metalogical entailment	252
<b>12</b>	<b>The problem of putative meaning and the logic of meaninglessness</b>	<b>256</b>
12.1	The problem of putative meaning	260
12.2	The delusion of meaningfulness	262
12.3	The logic of meaninglessness	265
12.4	Reflections on logics of meaning and meaninglessness	276
12.5	The logical priority of the bonded pair, reference-and-meaning, over truth-functionality	277

## DETAILED TABLE OF CONTENTS

<b>13</b>	<b>Projection</b>	<b>280</b>
13.1	A therapy for concepts	281
13.2	The term ‘projection’	284
13.3	Projection in relation to other forms of meaninglessness	287
13.4	Historical intimations of the concept of projection	288
13.5	The concept of projection	297
<b>14</b>	<b>Horizons</b>	<b>302</b>
14.1	Metalogical horizons	310
14.2	Limits of reference and boundaries of possibility and meaning	310
14.3	The detection of projections	313
14.4	Heuristics and the detection of horizons	317
14.5	Reflections on horizons	319
<b>15</b>	<b>De-projection</b>	<b>323</b>
15.1	The heuristic stages of de-projection	327
15.2	The epistemological neutrality and tautological nature of de-projection	330
15.3	Applying the method of de-projection	333
15.4	De-projection and framework self-enclosure	335
15.5	<i>Reprise économique</i>	338
<b>16</b>	<b>Self-validation</b>	<b>341</b>
16.1	Philosophical routes to certainty	341
16.2	Distant connections with self-validation	343
16.3	The “Worm of Ouroboros” logic of self- validation	346
16.4	The irrefutability of self-validation	350
<b>17</b>	<b>Rationality: Rules of admissibility</b>	<b>353</b>
17.1	The two sides of rationality	353
17.2	Intelligibility and coherence	355
17.3	Epistemological rationality	356

## DETAILED TABLE OF CONTENTS

### **PART III** **PHILOSOPHICAL APPLICATIONS OF THE** **METALOGIC OF REFERENCE**

	<b>Major Problems and Questions of Philosophy and the Philosophy of Science</b>	<b>363</b>
	<b>Methodological Recapitulation</b>	<b>365</b>
<b>18</b>	<b>Ontology and the metalogic of reference</b>	<b>369</b>
	18.1 The need for philosophical perspective	369
	18.2 The ontology of non-existent things	370
	18.3 Towards a general ontology of objects	373
<b>19</b>	<b>Discovery or invention in general problem-solving, mathematics, and physics</b>	<b>376</b>
	19.1 Mathematics and physics from the standpoint of the general theory of problem-solving	377
	19.2 Discovery or invention in mathematics	381
	19.3 Discovery or invention in physics	389
	19.4 Discovery or invention, according to Einstein	393
<b>20</b>	<b>The conceptually unreachable: “<i>The far side</i>”</b>	<b>395</b>
	20.1. “The conceptually unreachable” and the concept of horizon	396
	20.2 The finitude of what anyone knows and the finitude of the totality of human knowledge	397
	20.3 The incompleteness of knowledge	399
	20.4 The unlimitedness of our ignorance	400
	20.5 Thinking beyond the limits of thought	400
	20.6 Expressing the inexpressible: Reaching beyond the limits of language	403
	20.7 Fitch’s Theorem	405
	20.8 What <i>can</i> ‘non-omniscience’ and ‘knowability’ mean?	408
	20.9 Fitch’s Theorem: The de-projective result	408
	20.10 Unanswerable questions: Erotetic intractability	409
	20.11 The philosophical and psychological compulsion to trespass beyond horizon boundaries	413

## DETAILED TABLE OF CONTENTS

20.12	The need to resort to suggestive—e.g., rule-based or property-based—reference	415
20.13	Unbounded yet topologically enclosed frameworks	416
20.14	Conclusion	420
<b>21</b>	<b>The projections of the external world, things-in-themselves, other minds, realism, and idealism</b>	<b>421</b>
21.1	The projection of the external world	422
	1. The external spatial world	423
	2. The external ontological world	425
21.2	The projection of things-in-themselves	428
21.3	The projection of other minds	431
	1. External others	432
	2. Other minds as independent existents	432
	3. Other minds as things-in-themselves	433
21.4	The de-projective understanding of the external world and of other minds: The need to relinquish these projections	434
21.5	The projections of realism and idealism	436
	1. Realism	438
	2. Idealism	439
<b>22</b>	<b>The projections of time, space, and space-time</b>	<b>442</b>
22.1	The temporal basis of reference	442
22.2	Projections of time	444
	1. Projections of the past	444
	2. Projections of time-flow	444
	3. Projections of the future	445
	4. Projections of absolute time	446
	5. Projections of temporal constitutive subjective activity	447
22.3	The nature of spatial order	447
22.4	Projections of space	449
	1. Projections of absolute space	449
	2. Projections of spatial constitutive subjective activity	452
22.5	Projections of temporal and spatial continuity	453
22.6	Projections of space-time	454
22.7	In retrospect	455

## DETAILED TABLE OF CONTENTS

<b>23</b>	<b>The projections of causality, determinism, and free will</b>	<b>459</b>
23.1	Causality	459
1.	The conventional notion of causality	460
2.	The experiential basis of the conventional notion of causality	461
3.	The conventional notion of causality considered philosophically	462
4.	The temporal nature of causal relations	462
5.	Similarity relations between cause and effect	463
6.	The belief that the relation between cause and effect is one of “necessity” or “indispensability”	464
7.	The relation between cause and effect involves some kind of “tie,” “agency,” or the transmission of “productive power”	466
8.	The belief that the past regularity and uniformity of a cause-and-effect relation can be relied upon in the future	468
23.2	The evolution of the concept of causality	471
23.3	The functional understanding of causality	473
23.4	The concept of causal network	477
23.5	Determinism	479
23.6	Free will	481
23.7	Causality, determinism, and free will—in retrospect	485
<b>24</b>	<b>Projections of the self and of solipsism</b>	<b>489</b>
24.1	The self	489
24.2	Thinking entails a thinker—that is, there are always two components of thought: a subject (the self) and objects of which the subject is aware	490
24.3	Temporal preconditions of reflective reference: An aside	492
24.4	Projections of reflection and projections that thinking entails a thinker	493
24.5	The self is an existing entity	494
24.6	The self is the center of experience of the world	495
24.7	The self is a bearer or owner of its states	497



## DETAILED TABLE OF CONTENTS

24.8	The self possesses or is characterized by faculties which it exercises in a wide variety of ways	498
24.9	The self is an agent, the cause of thinking—that is, the processes of consciousness result from the activity of the self	499
24.10	Every experience is had by a self, by an at least implicit “I”—that is, consciousness by the self is a universal characteristic of experience	501
24.11	The sum total of the self’s experience comprises what metaphorically may be called a ‘container’ or a ‘phenomenal field’—that is, the self’s consciousness holds or encloses all that it experiences	502
24.12	The self serves as a limit which the individual cannot ever go beyond or get away from: The position of solipsism	503
24.13	The neutral character of pre-reflective experience	505
<b>25</b>	<b>Non-relational, agentless reference and referential fields</b>	<b>511</b>
25.1	Non-relational, agentless reference	512
25.2	Referential fields	514
25.3	Referential fields and the place of the observer	518
25.4	Referential fields and metalogical horizons: A brief summation	522
	<b>Relativity physics and quantum theory: Preamble</b>	<b>523</b>
<b>26</b>	<b>Relativity physics as seen through the lens of the metalogic of reference</b>	<b>525</b>
26.1	Introductory comments	525
26.2	The problem of correlating phenomena from the standpoint of different reference frames	528
26.3	Fundamental properties of the special and general theories	532
26.4	Steps from Newtonian physics to general relativity	536
26.5	The starting point: The Galilean transformation	538
26.6	The first step: The Lorentz transformation	539

## DETAILED TABLE OF CONTENTS

26.7	The second step: The shift to intrinsic identification	545
26.8	Einstein's mollusks and intrinsic reference	548
26.9	The concepts of tensor and tensor field	551
26.10	Invariance, covariance, and the metalogic of reference	556
26.11	The convergence of relativity physics and the metalogic of reference	559
<b>27</b>	<b>Quantum theory as seen through the lens of the metalogic of reference</b>	<b>567</b>
27.1	Introductory comments	567
27.2	Measurement-based perturbation	568
27.3	Eliminating "interpretation" from the Copenhagen interpretation	571
27.4	The Einstein, Podolsky, Rosen (EPR) position	574
27.5	Hidden variable proposals	578
27.6	Quantum theory and the projection of "underlying reality"	580
27.7	Indeterminacy and uncertainty	583
27.8	Complementarity: " <i>Contraria sunt complementa</i> "	586
27.9	The inseparability of the observer and the observed	589
27.10	The "collapse" of the Schrödinger wave function	593
27.11	The non-relational ontology of quantum phenomena	596
27.12	The reality of quantum discontinuity	600
27.13	Quantum theory as a model of objectivity	602
27.14	Quantum theory as a set of limitative results	609
<b>28</b>	<b>Epistemological lessons learned from and applicable to relativity physics and quantum theory</b>	<b>612</b>
28.1	The relation of mutual confirmation	615
28.2	The main epistemological lessons learned from and applicable to relativity physics	617
28.3	The main epistemological lessons learned from and applicable to quantum theory	619
28.4	The main epistemological lessons to be learned from relativity theory, quantum theory, and the metalogic of reference	623

## DETAILED TABLE OF CONTENTS

28.5	The wider applicability of these epistemological lessons	625
------	--	-----

### **PART IV HORIZONS** **629**

#### **29 Beyond belief** **631**

29.1	The negative science of the metalogic of reference	631
29.2	The metalogic of reference as a theory of error analysis and correction	636
29.3	Eliminative psychology and projective delusion	639
29.4	Obstacles to attempts to de-project belief	642
29.5	From conceptual therapy to disorders of thought: The human will to reach beyond its grasp	647
29.6	The psychology of projective belief	651
29.7	“Beyond belief”	654

#### **30 Critique of Impure Reason: Its results in retrospect** **657**

30.1	The negative science of the <i>Critique of Impure Reason</i>	657
30.2	The positive value of negative results	658
30.3	Delineating reality in silhouette	660
30.4	Intrinsic limitations of reference and identity	663
30.5	The <i>Critique of Impure Reason</i> and conceptual revolution	665
30.6	A possible future of philosophy	668
30.7	Intellectual humility: Submission of philosophers to the norms of science	670
30.8	The logical standing of the method and results of the metalogic of reference	672

### **SUPPLEMENT**

#### **The formal structure of the metalogic of reference** **675**

§1.	The advantages and the shortcomings of formalization	676
-----	--	-----

## DETAILED TABLE OF CONTENTS

1.1. The advantages	677
1.2. The shortcomings	678
§2. Internal limitations of formalization	680
2.1. The impossibility of comprehensive or total formalization	682
2.2. The impossibility of the total reflexivity of formalized systems	685
2.3. Formalization, temporality, recursion, and the metalogic of reference	685
§3. The formalization of complex or even ordinary reasoning	690
§4. The formal structure of the heuristic of the metalogic of reference	695
§5. Informal preliminary	697
§6. Deductive representation	702
§7. De-projection and the ontology of MoR: Informal comments	717
§8. Formalization continued	719
§9. Schematic summary of formalization found in this Supplement	724
§10. A concluding reminder	725

### APPENDIX I

#### **The concept of horizon in the work of other philosophers** 727

§1. The phenomenological horizon	728
§2. The horizon of transcendental Thomists	730

### APPENDIX II

#### **Epistemological Intelligence** 732

§1. Two approaches to the study of epistemology	732
§2. A set of epistemological skills	734
§3. From a defined set of skills to the recognition of a new variety of intelligence	737
§4. Is epistemological intelligence no more than a theoretical construct?	738
§5. The psychology of philosophers	741

## DETAILED TABLE OF CONTENTS

§6. The personality structure of philosophers as seen through the lens of confirmation bias	745
§7. Non-philosophical studies of so-called ‘epistemological understanding’	747
§8. Epistemological intelligence and individual differences	752
§9. Can epistemological intelligence be taught? Can it be learned?	759
§10. The projection of transcendence	762
<i>References</i>	765
<i>Index</i>	805
<i>About the author</i>	845



## Preface

This study is a descendant of research I began in the mid-1960s. At that time, Gabriel Marcel took a personal interest in the subject of my proposed dissertation and generously arranged with Paul Ricoeur for him to direct my doctoral work at the Université de Paris.

After I completed that work and had had my first taste of university teaching, during the academic year 1974-75 I was offered the opportunity to serve as research fellow at the Max-Planck-Institut zur Erforschung der Lebensbedingungen der wissenschaftlich-technischen Welt [Max-Planck-Institute for the Study of the Living Conditions of the Scientific and Technical World] in Starnberg, (then West) Germany. The Institut's staff at the time was quite small, consisting of its Director, well-known German physicist and philosopher Carl Friedrich von Weizsäcker (1912–2007), and theoretical physicists Michael Drieschner, Lutz Castell, and Hans Zucker. It was a special place, conducive to thought and writing.

During my fellowship there I had the opportunity to transform the phenomenologically based approach developed in my dissertation into what I believed, and continue to believe, can serve as a conceptually clearer, more precise, and less terminologically top-heavy approach to rigorous epistemology, one able to benefit by the tools of theory of reference and mathematical logic.

My work at the Max-Planck-Institut resulted in a monograph, *Metalogic of Reference: A Study in the Foundations of Possibility* (Bartlett, 1975), which was distributed within the Max-Planck-Institut as an in-house publication, and which therefore had a very limited circulation. My intention to fulfill my plans for that study, which I regarded as unfinished, was compromised over the years both by the obligations of university teaching and also by my own susceptibility to seduction to engage in research and to publish in other areas.

For nearly five decades, a persistently pestering monkey has managed to maintain a secure hold on my back, reminding me of that study's need to be completed, or at least in great measure completed since there is a great deal more that plausibly could have been included in this long book. Thanks to that annoying monkey, nearly half a century after my fellowship at the Max-Planck-Institut, I decided to return to what I originally called 'the metalogic of reference'. The present work has been strongly influenced by the earlier

study, I hope profiting by it, but extending it in ways I could not, when in my late 20s, have anticipated.

That an author can be “haunted” by an unfinished writing project is, for most people, an unfamiliar experience, one about which I will say a few words. To be haunted in the way I have been is to be reminded—regularly, stubbornly, persistently, and naggingly, sometimes on the periphery of consciousness, but often in a manner central to my awareness—that certain work needs yet to be done which, if left uncompleted, would on the day of my death leave me in a state of regret. Once this book has been finished, the monkey on my back will be free to slip away and return to the jungle and leave me in peace. As time has become ever shorter, I can no longer procrastinate by wondering whether I am perhaps not yet ready or not yet fully able to put my shoulder to the task. Once this book has been finished, I will be able to think and do other things freed of a life-long burden. With the completion of this study, I shall have brought to a close a project foreseen with considerable clarity in my late teens.

As I add these last few words to this work, I am convinced—and here I take the liberty of placing the dignity of humility momentarily aside—that this work provides compelling solutions to many of the main problems that have preoccupied philosophers for millennia. These solutions are strongly compelling in a special sense that is made clear in the course of the book: As we shall see, these solutions comprise results that cannot *not* be accepted without undermining the very possibility of meaning.

Perhaps few authors share Aldous Huxley’s reason for writing: As he expressed this, “My chief motive in writing has been the desire to express a point of view.... I write for myself and not for my readers.... [M]y dominant motive in writing is to make things clear to myself, and my writing is important to others in so far as it helps them to become clearer.”<sup>1</sup> Although such an unapologetic admission could be taken as a criticism of an author’s primary motivation, it has another side as an honest statement: that solving a certain set of problems has, for some authors, far greater importance to them, in and of itself, than service to unseen readers. So this has been for me. Notwithstanding this disclosure of my priorities, I have in the following pages made every attempt within my power to communicate to the reader in clear and unambiguous terms.

On my departure from the Max-Planck-Institut, Dr. von Weizsäcker shared with me his written thoughts relating to my monograph. Had I chosen to publish it, his reflections would have comprised that work’s Foreword. Since the method and soul of the older work live on in this study, in both the

---

<sup>1</sup> Quoted by Sullivan (1934, p. 141).



present book's fundamental approach as well as in several parts of its content, there is no better-qualified thinker to contribute a relevant Foreword than Dr. von Weizsäcker, whose comments follow.



## Foreword

by

Carl Friedrich von Weizsäcker

Director, Max-Planck-Institut zur Erforschung der Lebensbedingungen  
der wissenschaftlich-technischen Welt  
Starnberg, Germany

[This work's] goal is of a unique and difficult species: Dr. Bartlett seeks to develop a formal logical calculus on the basis of transcendental philosophical arguments; in fact, he hopes that this calculus will be the formal expression of the transcendental foundation of knowledge. He is certainly well-equipped for this task: his doctor's thesis showed a thorough familiarity with Husserl's transcendental philosophy (and also with Kant and Wittgenstein) and his mastery of modern mathematical logic and semantics is everywhere in evidence in the present book. He is, furthermore, sufficiently well-versed in the history of Greek logic to be able to locate his efforts against the background of the ancient beginnings in the Stoic-Diodorean and the Megaric-Aristotelian schools.

Dr. Bartlett is right, in my opinion, in developing his calculus from a very general notion of "possibility" that is limited only by restrictions which appear to be necessary preconditions of any theory. From these restrictions he develops the set of axioms that define his "metalogue of reference." The axioms are thus required to be "self-validating" in that their denial would result in a referential inconsistency; in this step lies the formalization of the transcendental principle. He is fully aware of the essential difficulty that his metalogue must be "self-referential," i.e., that the formal structure must explicate its own transcendental foundation. His careful studies of the scant literature on self-reference (R. M. Martin, B. van Fraassen) no doubt helped him to avoid pitfalls, but the construction of the calculus is entirely his own.

Dr. Bartlett sees the calculus as the carrying out of Kant's program of a "*phaenomenologia generalis*" (letter to Lambert, 1770)—a "negative science"—that was meant to precede any future metaphysics; his metalogue tells what referential statements are forbidden, everything else is allowed. I find it particularly attractive that his logic is ontologically "open," i.e. that the exis-

tential-referential commitments which he unavoidably makes actually forbid (as referentially inconsistent) the denial of existence to any sort of object. In this point his calculus agrees with Husserl's technique of excluding all apriori ontological characterization of objects, which is made to depend on an analysis of their modes of givenness instead. Another attractive feature lies in the referential inconsistency, entailed by the calculus, of all terms that imply a sharp subject-object decision (such as the term "observer-independent").

## Acknowledgments

Since the ideas and the method presented in this book had their beginning many years ago, acknowledgments of the support and assistance of individuals and institutions also stretch over many years.

In my first years of graduate school, and for the first time, I tasted the strong and disappointing flavor of self-limiting partisanship in philosophy, of blinkering fashions, restrictive paradigms, and preferred beliefs that so often stand in the way of the development of new and potentially revolutionary approaches—whether in philosophy or in other disciplines. Since acknowledgments can be both positive and negative, I mention this experience to acknowledge its eventual positive value and usefulness to my work by pointing me in promising directions. The need to find a dissertation director with a suitable background, interests, and openness to new ideas led me to contact a number of philosophers whose encouragement and willingness to serve as director of my dissertation I want to acknowledge. They included, in no particular order, Marjorie Greene, Maurice Natanson, Edward Teller, P. F. Strawson, Herbert Marcuse, and Paul Ricoeur, with whom I decided to work. All of these individuals expressed interest in a young doctoral student's proposal to develop a certain set of philosophically radical ideas and accompanying methodology.

Paul Ricoeur turned out to be a serendipitous choice, both because of his receptivity to new ideas and the fact that he was secure enough in his own thinking not to be intimidated by what is new and revolutionary. The other members of my doctoral committee, philosopher and mathematical logician Jean Ladrière and phenomenologist Alphonse de Waehrens, both professors of the Institut Catholique de Louvain, gave me, in their own respective ways, further much-appreciated encouragement. These men were intellectually open and provided their support to a doctoral dissertation that formulated an original epistemological method rather than the usual study of another philosopher's thought.

In the years following my doctoral work, I have benefitted greatly by the interest and intellectual sharing of a large and varied group of philosophers, scholars, logicians, mathematicians, and physicists. They are too numerous to list here, but I most especially wish to single out Robert Maynard Hutchins, Frederic Brenton Fitch, Henry W. Johnstone, Jr., and C. F. von Weizsäcker.

To institutions and relevant individuals, I wish to express my thanks to

the philosophy and mathematics faculties of Saint Louis University for their support and, I'm sure, their sometimes charitable interest in a strange and new approach which they perhaps may only dimly have perceived. Their tolerance for my unconventional and non-traditional thought made it possible for me to teach a wide variety of courses, from philosophy of science and philosophy of mathematics to continental thought to analysis, different levels of mathematical logic, computer programming, as well as the most popular course I have developed, a campus-wide course in quantitative problem-solving, supported by grants from the Lilly Endowment, the National Science Foundation, and the American Association for the Advancement of Science. The course was unusual in that it was accepted by the University as fulfilling a requirement for nearly all majors. In particular at Saint Louis University, I wish to thank philosopher of science Richard J. Blackwell and mathematician Ray Freese.

Later, by providing me with honorary faculty positions, my research has received the support of the departments of philosophy and psychology of Willamette University, thanks to the interest in my work and the initiative of Richard Lord and the backing of Mary Ann Youngren and Sally Markowitz, Chairs of the Departments of Philosophy and Psychology, and with thanks as well to the support of the Department of Philosophy of Oregon State University due to the interest taken in my research by its chairman, Peter List.

In connection with the earlier study that eventually has led to the present book, I wish to express my warm appreciation especially to physicist and philosopher Dr. C. F. von Weizsäcker and to physicists Drs. Hans Zucker, Michael Drieschner, and Lutz Castell, whose interest and support of my research made possible my visit with them, mentioned in the Preface, at the Max-Planck-Institut zur Erforschung der Lebensbedingungen der wissenschaftlich-technischen Welt in Starnberg, Germany. Thanks are also due to the Max-Planck-Gesellschaft for providing the funds necessary to carry out that earlier study.

Finally, I would like to acknowledge in advance the wished-for role of future readers: In a book largely devoted to the detection and elimination of conceptual error, it would of course be self-referentially unwelcome if some errors have managed to slip by the author. Should such artifacts of almost inevitable human fallibility be found, it is my hope that, by employing the very methodology developed in this study, the enabled reader will correct them.

## **Avant-propos: A Philosopher's Rallying Call**

**E**ven though philosophy has been called ‘the mother of the sciences’, even though the roots of rigorous mathematical proof are firmly embedded in the soil of philosophical antiquity, philosophy has frequently shrugged off these constructive and productive contributions of its past and has instead fostered a general Socratic-Kantian tendency to dissociate itself from strict demonstration and to embrace fondly the maxim which claims that one cannot learn philosophy but only to philosophize. After more than twenty centuries, philosophy can offer relatively few definitive solutions to philosophical issues of contention, and few solutions to broader theoretical questions. Proliferation of problems and of never-ending refinements in sophistication of statement and terminology indicates that the activity of the discipline continues on a strange, perpetually inconclusive course throughout a very long history. As a professional group, philosophers are often judged by the surrounding society to be academic throwbacks to misused aristocratic leisure, offering little of relevance to society’s utility-based interests. The very questions that occupy us have even been seen by some philosophers themselves as in need of linguistic therapy, while many philosophers tend, perhaps somewhat unconsciously, to accept that our principal function in the universities is to serve as an endlessly replayed recording of what others in tradition have written, and more often than not to show how much unsettled controversy has arisen in the dust of passing thinkers.

Judgments and pronouncements like these can of course be misleading and shallow, and yet they do point a shaming finger. No one knows how long intellectual vagueness and lack of focused orientation must fog the prevailing conceptual space before a discipline systematically achieves a place on which to stand. The lessons of the history of human culture are difficult lessons in patience.

Fortunately, however, the human species is by nature impatient. Two and a half thousand years call out to the philosopher’s mortality for less patience and more fruitful results.





## Introduction

Unless the scientist publishes the results of his researches so that they are accessible and can contribute to the general growth of knowledge his labors are ineffective. From the point of view of society they might as well have not been made at all, and from the point of view of the individual they are merely a form of self-indulgence. If ever use of the word ‘duty’ is justified, and we can certainly get along without it, I would be tempted to use it in this connection.

– P. W. Bridgman (1959, p. 291)

**T**he Roman poet Horace recommended that a serious author should withhold publication for nine years. Copernicus waited 36 years, four times that length of time. Here, I’ve postponed publication of a set of ideas five times longer than Horace’s recommendation, long enough that remaining life may not allow a greater multiple.

It is not often that an author has the opportunity—much less the motivation—to return to work undertaken nearly five decades ago, to review it with a more critical eye borne of experience and one would hope improved mental development, and then to weigh the pros and cons of resuming that work, substantially revising and extending it. More importantly, few projects initiated years ago realistically *merit* such work by their authors. The decision to do this has not been easy.

For one thing, a great deal has changed since the seed for this book was planted in 1974. Although the younger work was written from the standpoint of philosophy of science and mathematical logic, its implicit frame of reference was a combination of epistemology, theory of reference, and mathematical logic, with perhaps a mildly perceptible undercurrent of psychology. These different specialties are seldom combined, and when they have been, the result—in light of the literature published since the mid-1970s—has largely been to move in a direction away from the present book’s scope of interest and fundamental intent. There are, however, as we shall see, some signs that the trend of philosophical fashion and taste has nonetheless begun to change.

The earlier work from which this study developed was also influenced on

a fundamental level by a transcendental approach to philosophical investigation, and clear and strong signs of that continuing influence will be evident to readers in the pages that follow. Where Kant's *Critique of Pure Reason* sought to identify, describe, and analyze the preconditions of experience and knowledge, the present *Critique of Impure Reason* seeks—both in contrast to and yet in parallel with Kant's work written nearly two and a half centuries ago—to identify, describe, and analyze the preconditions of all referring or identification, and in so doing, to determine the universe of possible meaning. I have called this book 'critique of *impure* reason' because its purpose is explicitly *negative*: to recognize the boundaries of what is referentially *forbidden*—that is, to study the *limits* beyond which reference necessarily becomes devoid of meaning. As we shall find, attempts to transgress those boundaries, which will later be called '*metalogical horizons*', are both frequent and widespread. Such attempts comprise a broad class of conceptual confusions that lie at the very heart of many philosophical problems. A clear understanding of such attempted transgressions provides, as this *Critique of Impure Reason* seeks to demonstrate, a solution to many of these problems, a solution which rationally cannot *not* be accepted, as will gradually be made clear.

We shall find that the relation between the older and the newer *Critique* is conceptually basic and significant: We shall find in light of the conclusions reached in subsequent chapters that the *Critique of Impure Reason* possesses a logical and transcendental priority in relation to Kant's *Critique of Pure Reason*. Its priority in these two senses means that it comprises a necessary preliminary study, conceptually more basic because its tasks of error-detection, correction, revision of concepts, and, in some cases, their elimination, must precede the more constructive efforts of the *Critique of Pure Reason*—and indeed must precede the efforts of *any* coherent theory developed to account for the objects it wishes to study. A study of "impure reason" must, of necessity, have that precedence in order to insure that subsequent constructive tasks are not contaminated, handicapped, and even undermined by the major and unrecognized variety of error that is the central focus of the "negative science" of the *Critique of Impure Reason*.

By the time readers have reached the concluding chapters of this book, it will be clear both how and why this is necessarily so. At that time, we shall examine the concept of "negative science" and its critical function both in philosophy and in the analysis and appraisal of a wide range of commonsense concepts, claims, and beliefs, as well as their counterparts that are employed by natural and formal science. We shall find that many of these concepts, claims, and beliefs have provided a major, perhaps *the* major, subject-matter of philosophical controversy during the past two millennia.

Although this work bears the relations to Kant's work that I have briefly touched on, the present study is not a study of Kant's thought, nor is it developed on the same conceptual level, nor is its methodological approach the same, nor does it share many of the same concerns, nor does it accept many of the principal conclusions Kant reached. The path chosen in this book diverges, radically and in many basic, explicit, and important ways from Kant's, something which should come as no surprise, and indeed should be hoped for after the passage of more than two hundred years. Although readers will find occasional references and brief discussions of Kant's work, this study is not intended and should not be thought to be a contribution to Kant scholarship, a contribution to continuing Kant studies, or an attempt to advance Kant's own individual philosophical objectives.

Despite these disclosures and the significant divergence between the goals of Kant and the objectives of the present approach, the reflective reader will find that the "negative" and "positive" *Critiques* complement one another in certain ways—as conceptual models which, taken together, may serve as paradigms of interrelated, conceptually necessary approaches to a comprehensive philosophical understanding of reality.

I mentioned above that in the ancestor of this work there was perhaps a "mildly perceptible undercurrent of psychology." When these are relevant, occasional psychological observations occur in the present book, but they are few and far between; the book is by no means a work in psychology. Nevertheless, a few words about the connection with psychology may be of value: The application of psychology within a philosophical and sometimes epistemological context has resulted, principally during the past two decades, in a variety of books and papers that have made use of the sobriquets 'therapeutic philosophy' or 'philosophy as therapy'. Much of this work has followed in the shadow of Wittgenstein's focus on language and the claim that language is at times used inappropriately and specifically in philosophically "mystifying" ways. Some self-credentialed "philosophical therapists" have more recently sought ways to lend credibility to their professional acceptability as non-traditional, alternative "therapists" or "counselors" who "treat" real "clients" for real human problems, and a few professional organizations have been established to support and authorize them.

An exception to this development that sees in "therapeutic philosophy" the potential to help "clients" to cope with and perhaps to overcome some of their problems of daily living has been my own altogether different and unrelated work in a series of studies published both before and after my 1974 monograph, *Metalogic of Reference*. In this research and publications, beginning in the 1960s I coined the terms '*conceptual therapy*' and '*conceptual*

*pathology*', neither of which had anything to do with "treating clients" in an alternative and contrived semi-clinical setting.

Instead, these terms, as I defined them, refer to a form of conceptual analysis whose "therapeutic purpose" is to identify, correct, replace, and avoid faulty, self-defeating *concepts* (and not misused or beguiling *language*). These are concepts that cannot, *in principle*, serve their intended functions due to the fact that they can be seen to be self-referentially self-destructive from the standpoint of an abstract meta-level which we shall later call 'maximum theoretical generality'. These ideas are developed in some detail in the present work and in others of my publications; I mention them in general terms here to underscore their remoteness and divergence from "client-centered philosophical counseling."

While the philosophical counseling of real people in an allegedly clinical setting has attracted a few philosophers, mainstream philosophy today has continued a general movement that during the 1960s began to veer decidedly away from a proof-oriented interest in establishing permanent and unimpeachable results. Since approximately that time, philosophers have largely given up such a goal, and doubts have come to dominate the profession that such a purpose is genuinely realistic, or that it can, in principle, be realized, or even that it *should* define an appropriate purpose for philosophy. Conceptual relativism has come to dominate much philosophical discourse and study, buoyed by anthropology's powerfully influential recognition of cultural relativism and by the endorsement of relativism in a society obsessed by strictures of political correctness and the consequent legitimization of relativist values. At the same time, language analysis has attracted much of the attention of Anglo-American philosophers, while structuralist, post-structuralist, hermeneutic, deconstructionist, modernist, post-modernist, feminist, narrative-oriented studies (among others) have come to dominate the thought and literature of European philosophy, and to define the interests of most of the remaining population of the English-speaking world of philosophy.

Since the millennium, a number of addresses have been presented before the American Philosophical Association that have underscored—through negative criticism as well as subdued praise—the fact that philosophy has, in its more than two thousand year history, virtually never (perhaps just plain 'never' is the honest and accurate judgment) produced *any* results which are widely accepted, which are recognized as firmly demonstrated, which are resistant to contention and controversy, and which can constructively be built upon by future generations of philosophers so as to create a body of conclusions that represent clear and irrefutable results produced by the mental labors

of capable, skilled thinkers. Here are samples of observations presented before the APA:

What can progress in philosophy be, if it is compatible with so much ineliminable disagreement concerning fundamental issues? (MacIntyre, 2010, p. 70)

[I]f it is true that, 2,400 years after Socrates, we have not come up with a single successful argument for any substantive philosophical thesis, it seems to me that that should generate at least a bit of a worry about our discipline. (Tooley, 2011, p. 30)

Most philosophy..., I believe, is such that it would have been no loss to the world if it had never been published. (Wolf, 2011, p. 47)

[I]f one asks a philosopher for even a *single* book that will summarize the elements of philosophical knowledge—as one might ask a chemist for a handbook of chemistry—he will have nothing to present. There *is* no general, agreed body of philosophical knowledge.... [I]f we examine the history of modern philosophy, it appears to be *a subject in search of a subject matter*....

This should give us pause. How can it be that after two and a half thousand years of endeavour philosophy has still not reached the status of a science, has no agreed subject matter, and has no fund of philosophical knowledge? How is the poverty of philosophy, construed as a cognitive discipline, to be explained?...

The promise that after two thousand years of irresponsible adolescence, philosophy will at last produce a flood of truths and well-founded theories—tomorrow—has been made, and proven empty, far too often to carry conviction. (Hacker, 2009, pp. 130-131, 134)

The idea that there are proofs in philosophy as there are proofs in mathematics is ridiculous, or not far short of it.... Only one thing can be said against this standard of philosophical success: if it were accepted, almost no argument of any

substantive philosophical thesis would count as a success. (van Inwagen, 2006, pp. 37-38)

The aspirations of philosophers of the past to transform philosophy into a decent scientific discipline, when collected in a historical survey, now seem to us nothing but a boulevard of broken dreams. (Philipse, 2009, p. 163)

The ideal of philosophy as a would-be rigorous science has become, if not laughable among the majority of philosophers, then for them a mirage to be waved aside in weary or condescending dismissal. But yet there still are—here and there—as some of the sample quotations above intimate, signs of growing dissatisfaction. One of the most surprising of these signs has been expressed not very long ago by occupants of the Wykeham Chair of Logic:

[T]he rhetoric of finally founding philosophy as a rigorous theoretical enterprise has become popular in Oxford quite recently, where it is used by occupants of Oxford's Wykeham Chair of Logic. Timothy Williamson, for example, recently has urged the need for rigorous methodological standards for philosophy, and has called those who oppose “systematic philosophical theorising” as succumbing to an “unnecessary surrender to despair, philistinism, cowardice or indolence.” (Philipse, 2009, p. 162)<sup>2</sup>

There have been a few indications like these that “philosophy as rigorous science” has not entirely been relegated to oblivion. Perhaps the pendulum is beginning, weakly and hesitatingly, to change its direction of swing. But perhaps not.

It continues to be a challenge to discern any emerging pattern as we scrutinize the uninformative tea leaves. The overwhelming general consensus that has taken shape and fossilized during the past four to five decades asserts two propositions: (1) philosophy has, during its long history, *in fact* never established any results of a high degree of reliability and unquestionability; and (2) philosophy has either (a) *failed* to achieve such successes, or (b) it is *a mistake* to criticize the discipline for not accomplishing what it should never have been expected to accomplish. The first part of this general consensus is empirically based, decided in the face of the simple and undeniable evidence

---

<sup>2</sup> Williamson, incidentally, goes so far as to reject the language analyst's mantra that the primary task of philosophy is to construct a systematic theory of language.

that no examples of philosophical results of the rigorous indisputable kind can be produced. The second part of the consensus hinges upon matters of value: If one values firmly demonstrated results, then anyone who adheres to perception (2a) is destined to be disappointed and should probably move to another field of research—say, science or mathematics. Alternatively, if he or she sees things according to (2b), then all is well in the current multicultural, diversity-affirming, and inconclusive universe of philosophical dialogue, argument, and contention.

Assertions (2a) and (2b) express personal professional decisions as to how one should spend one's time and labor. To confront the *de facto* absence of provable philosophical results with any sense of realistic optimism that such results may yet be forthcoming, after such a long past that extends through millennia with no solid conclusions to show for the effort, requires a very considerable degree of hopefulness, determination, and willingness to oppose the fashion and style that define philosophy today. Nevertheless, from the point of view of this book's author, having been born with a strong stubbornness of character, and having from a young age been a person for whom conformity with prevailing disciplinary fashions means little, I have been willing, and have preferred, to set an independent course, as readers of this optimistic study will find.

It is a fundamental meta-truth that truth itself has no direct connection with popular consensus, and to equate the two is to stretch very considerably beyond its meaningful application democracy's unquestioning love for the shared beliefs of groups tallied by voting. And yet reliance upon group consensus has, especially in recent years, become a firmly rooted way of placing the crown of Truth on beliefs that happen to meet with social and disciplinary approval. And yet, as intellectual history tells us, individual, independent effort has often proved to be more promising in the search for demonstrable truth than the consensus of group beliefs.<sup>3</sup>

As readers turn the pages of the book that follows, it will be clear that this study runs against the prevailing grain in a number of fundamental ways. If it is successful in its own terms and in relation to its specific goals, then it may stand as a counterexample to the non-existence of demonstrable results in philosophy. And should it fail to establish such results, it would be inconsistent to fault the efforts made here, given the context of a discipline that continues on a perpetually inconclusive path through time.

If I am sometimes assertive, and perhaps for some readers seemingly dogmatic, it is in the interests of economy of presentation. I have generally

---

<sup>3</sup> For readers interested in convincing, historically based evidence supporting this and related claims, see Murray (2003).

chosen in this work not to discuss in detail philosophical approaches that may comprise alternatives or opposition to what is presented because this has the unavoidable tendency to mire the discussion in controversy rather than to clear a way for constructive development. It has not been my purpose in this study to criticize the ideas of others, but rather to explain in detail how we can make steadfast philosophical progress that will advance both our knowledge and the long-lacking rigor of our discipline. In the chapters that form Part III of this book, readers will find an extensive series of applications that show how inquiry can be conducted based on the principles formulated in earlier chapters. Throughout, I have felt that a direct, clearly stated, assertive presentation provides the best approach given the goals of this study.



## A Note to the Reader

### *The level of difficulty of this work*

The subject-matter of this book is admittedly difficult. But much of its difficulty is not due to the intrinsic complexity and the intellectual demands of the subject-matter itself, but it is rather due to the degree to which that subject-matter is likely to be unfamiliar. The approach developed in this book is new; it breaks new ground, and does this in a new way. More than this, however, it runs counter to much of today's "mainstream philosophy," and, more significantly—and more challengingly—it requires of the reader a willingness to think in counterintuitive, counter-habitual, and counter-conventional ways.

To accomplish the objectives of a study of this pioneering kind has not been an easy task for its author, who has had to create an unfamiliar vocabulary in order to communicate unfamiliar concepts to readers, some of whom will no doubt find it hard, and may perhaps be unwilling, to suspend or to place in question their accustomed conceptual frameworks.

For readers who take pleasure in "thinking outside the box," the subject-matter will seem considerably less difficult and demanding. For other readers who may be less secure in embarking on an exacting, intellectually self-critical adventure, to be asked to think "outside the box" may provoke anxiety and, as a result, the subject-matter's difficulty may appear to be unduly magnified.

It is of course my hope that I have found a way to communicate effectively to both sorts of readers. But the ability of a book to communicate to the reader is never one-sided: There is no such thing as a book that in itself communicates well, taken out of relation to effective readers. And so I must also hope that readers with the necessary mental openness, interest, and skills find their way to this book.

### *The general absence of examples in Parts I and II of this book*

The first two parts of this book develop a method for solving certain epistemological problems. It is my conviction that until that method has been clearly formulated and then understood by readers, there is little point—along

the way—in providing examples of its application. Applications by means of examples that are introduced prematurely are more likely to mislead and confuse than they are to be enlightening. Only when the method in its entirety has been described will it make sense to attempt to apply it to specific theories, positions, and concepts.

This book's ideal reader is a rare individual: someone who can delay the intellectual gratification of confronting specific epistemological problems, who has a fair amount of patience in remaining on a highly abstract level of reasoning without a pressing need prematurely to apply that reasoning to concrete instances. Admittedly such ideal readers are scarce. For those readers who are not among these rarities, I recommend the cultivation of trust that the author has made a sincere effort both to communicate his meaning as clearly and as simply as he could, and to fulfill the intentions and promises made in this study. With a certain amount of patience and fortitude, the reader will be rewarded when it comes time to consider real problems.

*The length and cumulative nature of this book*

Due to the large number and variety of concepts, claims, positions, and theories which this study analyzes, this book is of necessity long. As a result of its length and the fact that the results reached are progressively cumulative, building upon one another, I have provided occasional brief recapitulations interspersed in the text in order to take stock of steps that have been made.

## A Note on Conventions

During the course of this book we shall encounter certain widespread human psychological dispositions to believe without adequate justification. One of these is to mistake words either for the things they represent or for the meanings they express. To avoid such misplaced belief it is essential to distinguish two fundamentally different uses of language.

To do this I use the well-established convention of semiotics, the theory of signs, to make clear when so-called ‘autonymous’ or ‘indirect reference’ is made to a word, phrase, or other symbol, and to distinguish this from its ordinary use. This convention is a reminder that we need to be aware when reference is only to words themselves, as opposed to what they mean or what we take them to designate. When reference is made, then, to a word, phrase, or symbol itself, single quotes (inverted commas) are placed around it. To illustrate: ‘one’ contains three letters. Single quotes are also used to set off a quote within a quote. Double quotes are reserved for direct quotations and to draw attention to words employed in an important, odd, exaggerated, or illogical way that extends or distorts their usual meaning.

Whenever feasible I’ve used gender-neutral language in this book. In infrequent passages where it would be excessively repetitious to use ‘he or she’ and its variants I’ve followed the equitable convention proposed by Charles Murray (2003, p. xiv) to use the author’s own sex as the choice of third-person singular pronouns.

Internal references within this book to chapters, chapter sections, and/or chapter sub-sections are indicated in braces: ‘{2}’ refers to Chapter 2, ‘{7.3.5}’ refers to Chapter 7, Section 3, Sub-section 5.



*A horizon defines, from your present standpoint,  
how far you can see.*



# PART I

## WHY PHILOSOPHY HAS MADE NO PROGRESS AND HOW IT CAN

[The] evident and insurmountable plurality [of philosophical positions] constitutes the central aberration of philosophy because it leads to the destruction of the very idea of truth in philosophy.

– Jérôme Grynepas (1961, p. 429; author’s translation)

During the past two and a half millennia, the central ideal and often restated objective of philosophy has been the pursuit of truth. During much of the 20th and so far in the 21st century this goal has largely been abandoned. In its place a relativist multiplicity of approaches has taken root that variously expresses the special interests of multiculturalists, feminists, philosophers of language, postmodernists, hermeneutical philosophers, post-structuralists, and others—many of whom urge that we focus our understanding on smaller, more detailed, and more readily studied subjects. The pursuit of detail by philosophers has frequently displaced the pursuit of *demonstrable truth*, which has largely been given up as an empty and mistaken ideal. The big picture of the place of the human species in the universe, of what we can know as opposed to what we merely believe—interests that defined the philosophical endeavor to reach the certainty of truth—have generally been discarded in favor of ever-smaller snapshots of such things as the ways in which we use certain words, how we interpret texts and “narratives,” and how we defend the interests of our genders and our social and political affiliations.

In the first two chapters of this book I put diplomacy aside and attempt to provide a realistic and unflinching answer to the question, Why has philosophy made no constructive, definite, conclusive progress during its long tenure in human minds? In seeking an answer to this question, my wish is to understand why it is that the ideal objective of philosophy to reach demonstrable truth has largely been given up by the great majority of philosophers. The approach that I employ in answering this question is psychological, for the limitations that have held back the discipline of philosophy have been and

continue to be, in the author's judgment, inherently psychological in nature.

For readers who do not share an interest in examining the philosophical enterprise in psychological terms, or who may be affronted by a direct, plain-spoken, overtly critical psychological assessment of the discipline, the first two chapters of this book can be skipped with no loss of its main philosophical content.

From the standpoint of this initial understanding of the internal psychological limitations that obstruct genuine progress in philosophy, the body of this book then proceeds to develop an approach to philosophy that does not suffer from the inherently questionable nature of a "philosophical view" based on one individual's or a group's set of *preferred beliefs*, but instead provides a constructive, definite, and conclusive basis we cannot *not* accept without fundamental and self-defeating incoherence.



# 1

## Philosophical-psychological Prelude

[T]he discussion of psychological questions is out of place in a philosophical enquiry.

– A. J. Ayer (1952/1936, pp. 121-122)

But what a mysterious taboo it is! Does it mean that, if an author promises to be strictly philosophical in the sense of introducing no psychology into his book, then he should keep his promise and introduce no psychology? But there seems no reason to make such a promise.... It is a psychological fact that the human mind thinks...; and the adequate investigation of this must be both psychological and logical.

– Richard Robinson (1962/1954, p. 14)

**D**uring approximately the past one hundred years, the attitudes, beliefs, and values that define majoritarian philosophy have undergone an unmistakable shift in preferred fashions: Some philosophers have dignified this as a “paradigm shift,” but fashion combined with fad it most fundamentally is. A recognition of this shift should increase our critical self-awareness. The changes in philosophy have been profound and affect what is taught, the type of students who are drawn to this teaching, and the type of teachers who do the teaching. The same changes affect what is written in the name of philosophy, what is published, and what is read. They affect, in a self-fulfilling manner, what we expect the discipline to accomplish and how it is to satisfy those expectations. The attitudes, beliefs, and values that define the boundaries of the framework of current philosophical thought are—like all human attitudes, beliefs, and values—self-limiting: They establish limitations of the acceptable and the preferred, and in doing this they erect barriers beyond which it is unacceptable to go.

The following table describes some of these recent changes in philosophy that have occurred over the past century:

Views which philosophy is moving away from, or is explicitly rejecting:	Views which philosophy is finding “more congenial,” or is explicitly advocating:
The belief that philosophy is capable of reaching an understanding of the world formulated on the basis of such intellectual criteria as conceptual clarity, distinctness, certainty, invariance, universality, etc.	The belief that philosophy must content itself with much more modest goals, embracing such attitudes as relativism, contextualism, multi-culturalism, etc.
A commitment to a broad and integrative focus that emphasizes unity of knowledge and the recognition of systematically inter-related connections in order to make possible the development of a total comprehension of reality.	A turn away from “system building” in favor of technically accessible, precisely delimited topics, studied with great attention to detail through the construction of examples and test cases.
A focus upon “the great problems”: the questions that are most fundamental and meaningful from the standpoint of understanding the human enterprises of science, mathematics, and the humanities, and of one’s life, relation to other people, and place in the overall scheme of things.	A focus upon increasingly “well-defined,” specific, and less “grandiose” topics, upon technical problems that are posed, not by life, but by other philosophers; hence the focus becomes discussion-oriented, in which what is discussed is the content of past discussions. In other words, sight is lost of the attempt to understand reality, and instead participants discuss and critique <i>each other’s views</i> .

<p>An interest that is subject-directed, committed to resolving questions which have broad significance and application, studied in a manner that directs philosophical attention to fundamental concepts, first principles, the premises of human thought.</p>	<p>The belief that philosophers of the past have been misled by a focus on broad conceptual issues undertaken by means of theoretically abstract conceptual analysis. It is therefore incumbent upon philosophy to take a much more concrete and down-to-earth approach, studying, e.g., the uses of words, phrases, and sentences, often coupled with an interest in analyzing “narratives” through textual analysis.</p>
<p>The belief that philosophical thought is capable of reaching a level of truth that is resistant to disqualification, doubt, contention, or opposing argument, and that truths which are attained in this way have enduring value, i.e., are “perennial.”</p>	<p>Interpretation of texts and of the content, form, and intent of narratives becomes the central concern, in the belief that such interpretations are human constructs that have only internal hermeneutic significance.</p>
<p>A high value is placed on attained authority, i.e., respect for exceptional originators of integrative insight who have been able to formulate their understanding in a unified and systematic way.</p>	<p>A style of anecdotal example-analysis, of frequent quotation and small-scale criticism is embraced; discourse becomes egalitarian so that the “playing field” is leveled; all players are equal because all concentrate on the same highly delimited topics. The technical jargon developed to discuss these topics imbues players with the confidence of expert specialization.</p>

<p>Related to the above respect for authority is respect for the demands—and, in particular, the “policing value”—of reason, rationality, logical argumentation, and demonstration by conceptual analysis.</p>	<p>The “policing function” of philosophy is dismissed in favor of an openness to alternatives, and an embrace of alternative “stories,” “narrative interpretations,” and textual explications.</p>
<p>There is the sometimes advocated claim that philosophy has the capacity to be “scientific” (e.g., by Reichenbach and in the early work of Husserl), asserting an in-principle-achievable philosophical goal in which past results can be built upon so that over time a unified, provably valid conceptual understanding of the world can be developed.</p>	<p>Philosophy should repudiate all attempts to become “scientific”: It is a misunderstanding of the philosophical enterprise to suggest that “definitive results” have a necessary place in philosophy. Instead, philosophy is “process”; it is an activity, not a search for demonstrable results; philosophy clarifies but does not answer.</p>

### 1.1 Table of paradigm changes in philosophy

There is an unmistakable psychology at work in this shift from the attitudes and claims found in the left column to those in the right. Consider the following two real examples:

After reading a philosophy colloquium paper at Saint Louis University, distinguished metaphysician Leonard Eslick was asked to give some examples of metaphysical results that have been firmly established as a result of more than two millennia of philosophical thought. Eslick hesitated for an uncomfortably long time—clearly uncomfortable both for him and for his audience—and then replied by stating that the proposition “Everything is unified” is one such definitive result.

There was, despite the Department’s sympathy and respect for Professor Eslick, an audible shifting of feet, creaking of chairs, pained silence, and group embarrassment. That more than two thousand years had led us to *this*—perhaps metaphysics’ most notable truth—was mortifying and an uncomfortable and dubious reflection on the value of metaphysics. Here, frustration and

disappointment probably characterized the psychology of many in the audience.

At another philosophy colloquium at the same university, distinguished philosopher Paul Arthur Schilpp read a paper followed by the usual question-and-answer session. He was asked, on the basis of his many decades devoted to the famous series of books edited by him, *The Library of Living Philosophers*, what he had learned in his close contact with many of the philosopher-luminaries of the twentieth century. His response was short and to the point: "Philosophers do not *want* to understand one another." He went on to explain: They are inherently (this is my paraphrase) narcissists in the realm of theory. Each advocates his or her own opinions, with little or no concern for what anyone else thinks, in the spirit of: "Your alleged facts are mere opinions; my opinions are unquestionable facts." Philosophy, as Schilpp intimated this, functions as an authorized political-like platform on which philosophers can stand, with the approval of society and of higher education (and be paid by them), so that philosophers may voice their own personal views, and do this in a fashion deaf to what other philosophers, standing on their own territorially protected platforms, happen to be shouting. Here, the psychology is of one of narcissism, the psychology of adherents to self-encapsulated ideologies who do not *want* to legitimate or communicate with one another.<sup>4</sup>

Philosophers are no less psychologically vulnerable and affected than the rest of us. Their psychological receptivity to classical ideals after centuries of disappointment—by those susceptible to such disappointment—has had an unmistakable effect on the direction philosophy has taken. The loss of psychological receptivity to the attitudes, beliefs, and values that appear on the left-hand side of the previous table is altogether understandable, and even, I hazard to say, predictable. Ideals that consistently have led to frustration and disappointment lead to a form of disenchantment characterized by attention to smaller and more manageable details of analysis, and often to ideological self-enclosure.

In the context of such prevalent disenchantment, any attempt to look with fresh eyes at the older ideals is bound to be greeted with skepticism if not outright rejection. No one wants to be led down yet another blind alley.

Pick virtually any book from the shelves, whether fiction or non-fiction, and if you will look for its telltale signs, you will find that its author seeks, in one guise or another, to persuade. Authorship is a disguised variety of salesmanship. And so, rhetoric, the art of persuasion, plays a fundamental psychological role in the majority of books.

This applies no less to the present book. For many years, I have struggled

---

<sup>4</sup> For more detailed discussion, see Bartlett (1986a).

with the question how to overcome the loss of psychological receptivity of today's philosophers to conclusive, methodologically strict, result-oriented philosophy. This unanswered question has in great measure been responsible for my reluctance to reconsider the basic approach formulated in a work originally written nearly half a century ago—a work which at the time and during the intervening years I have allowed prevailing philosophical fashion to discourage its further development and publication. During the intervening years, a significant number of which I have devoted to research in clinical psychology, rather than acquiring constructive insight into how philosophy's losses might be rectified, I have instead discovered layer upon layer of solidly confirmed psychological facts concerning the human constitution.

These facts have forced a realization on me that, if philosophers are to regain a psychological receptivity to and appreciation for rigorous philosophy, this is a matter over which a single author can exercise very little influence: Here lies, at a most basic level, the familiar human fact that the pendulum's swing is governed wholly by the momentum of the preferred views of the human majority. When, and if, the pendulum of fashion will swing back in a direction supportive of the goals of this work, are questions I of course cannot answer.

It has taken some decades for me to make a self-conscious choice to disregard prevailing tastes, beliefs, fashions, and fads—in other words, simply to ignore the current paradigms of philosophical research and to try *once more* in philosophy to reach for that which has alluded us for so long—those definitive, unarguable, provable, solid, constructive results that are immune to revision, that can stand the test of time, and that can be used as stepping stones for others who would make constructive, incremental progress in a discipline that has yet provided no such results that can be built upon. We recall these comments from three voices from the past:

“As the saying goes, may God preserve us from all that we can get accustomed to.” – Isaac Bashevis Singer (1982, p. 134)

“Man simply cannot accept human limitations as inevitable in the scheme of things.” – Ernest Becker (1975, p. 165)

“It is not at all necessary to hope in order to endeavor, nor to succeed in order to persevere” – Reputed to have been uttered courageously in the face of overwhelming odds by William of Orange. (Rüstov, 1980, p. xxix).

### 1.1 Bifurcations of the human mind: The rational bridge problem

Beyond the ingredients of the personal incentives and commitments of philosophers, there is another psychological factor to consider: the relationship between rationality and conviction. It has been my long experience with fellow human beings—and most especially fellow philosophers—that rationality and conviction seldom go hand-in-hand. In observing their underlying psychology, conviction that preferred beliefs are true *by virtue of the commitment invested in them* takes very considerable precedence over whether they are the consequences of rational thought.<sup>5</sup> Peter Suber, philosopher-attorney and now specialist in formulating open source publishing policy, has expressed this well:

Whether a reader will reject a premise or rule is contingent on that reader's individual psychology and prior convictions, and no amount of care in insuring the formal validity of the reasoning will make these contingencies irrelevant. I describe the conditions of unpersuasiveness in psychological terms in order to emphasize, not their relativity to individuals, but their independence from the formal properties of the argument studied by logicians.<sup>6</sup>

The psychologically normal human mind is *bifurcated*: On one side, there is an area reserved for desirable beliefs that bring various gratifications, and on the other side is an area in which rational thought occurs. Only in some individuals, who make up a relatively small minority, are these two areas of the mind *fused* and *intercommunicate* so that rational conclusions *determine* the individual's genuinely felt convictions and conduct.

Philosophers are no less subject to this bifurcation of the mind, and this psychological reality is what I wish here to focus attention on and underscore. The bifurcation in question I will call '*the rational bridge problem*', in parallel with what elsewhere I have termed '*the moral bridge problem*' (Bartlett 2005, 2011). Both bridges pose fundamentally important psychological problems for humanity. The moral bridge problem concerns the fact that there exist people who have well-developed moral *reasoning* skills, and yet those skills fail to carry over into their moral decision-making as it is expressed in

---

<sup>5</sup> Discussed in Bartlett (1969-70).

<sup>6</sup> Suber (1994, p. 245).

their actual *conduct*. It is this lack of “carry-over” that constitutes one of the most important of human moral concerns.

Considerable educational effort has now and again been made to develop in students the ability to reason their ways through complex moral dilemmas, and to formulate morally enlightened choices as a result. But there is no evidence that, once having acquired such moral reasoning skills, these students will *behave* any better than their morally untutored peers when it comes to the willingness of the great human majority, when circumstances are “right,” to engage in state-authorized aggression and killing in wars, participation in judicial executions, perpetration of school and adult bullying, domestic abuse, endorsement of torture in the name of national security, depredation of the world’s natural resources and biodiversity in the interests of human development and financial gain—a list that could be continued at some length. The moral bridge is a bridge that relatively few cross automatically and naturally, from morally reasoned judgment to moral conduct.

In the same way, relatively few people who are capable of high-level skills in their use of reason actually feel strong personal commitments to the conclusions to which their own rationality brings them.<sup>7</sup> To quote a passage from Bartlett (1969-70) which points to the importance of both the rational and the moral bridge problems:

“[F]or Socrates the essential characteristic of all the oral arguments which he pursues is not truth but *conviction*; the conviction which welds belief to action” (Sesonske, 1961, p. 45).

In this way, Sesonske ... draws attention to the Socratic concern that argument must go hand-in-hand with the conviction which leads from the realm of philosophic discourse to the world of practical action. The strength of the link between valid argument and commitment to its conclusions as a

---

<sup>7</sup> In my study of the psychology of normality and its numerous shortcomings, attention is drawn to the strong evidence that testifies to the small percentage of the human population that has genuinely good psychological health (cf. Bartlett, 2011, Appendix III, “The Distribution of Mental Health”). The even smaller percentage of this already very small group that has a psychological constitution that welds rationality to conviction gives one reason to restrain optimism concerning the degree to which people are willing to be receptive to and to act upon the results of rational investigation. Those who are capable of true intellectual and emotional openness to new views, which have the potential to displace their own preferred beliefs, are extremely few in number. The ability of people—for *it is* an ability—to listen dispassionately when confronted with sound reasons to adopt results that conflict with their favored beliefs, and then to form convictions consistent with this experience, is an ability that is extremely rare.



guide for action depends on the extent to which the argument has persuasive force and can convince us that what is true should be adopted by us in our everyday lives....

[The Socratic process of question and answer] is governed by the supposition that if truths are arrived at, those involved in dialogue will be *compelled* by the force of the argument to acknowledge these truths. Not only will they come to accept these truths, but with this acknowledgment will come a fundamental conviction that will over-flow the boundaries of argumentation and will influence them to live according to the truths which they have been persuaded to affirm.

It is clear that this connection between argument and practical affairs, between knowledge and action, between the conceptual affirmation of principles and influence upon daily behavior, can and does break down.....

Why not accept the breakdown and give up the attempt to persuade and convince?...

To Greek thinkers, reason is the very condition of man's ... moral being.... It follows that in the Greek conception, the moral life is practically identical with the rational life. The best life is the one lived according to the prescriptions of rationality: for it is the life in which action and thought are wedded as means to the end or *telos* of life, which Plato calls the Good....

A man's life is good in proportion as it exhibits a purpose which directs his action; the more a man's life assumes a structure and a plan by virtue of rational goals, the more his life is good. Man's life becomes intelligible and good in proportion as rationality serves to guide his thought and action.

Now the connection between rationality and life led according to its prescriptions is *conviction*. Conviction is the essential link between reason ... [and] moral action. When this link breaks, Plato's concern begins. His is a concern to maintain and to strengthen the bridge between the reasonable and the desirable; without this tie, morality is rendered impossible.<sup>8</sup>

---

<sup>8</sup> It is interesting in this connection to note Kant's view of 'conviction': "If the judgment is valid for everyone, provided only he is in possession of reason, its ground is objectively sufficient, and the holding of it to be true is entitled *conviction*" (Kant, 1965/1929, A820, B848, p. 645).

The ideal goal of the approach to philosophy that this book presupposes is that it is desirable, and indeed humanly important, to overcome the present bifurcated condition both of philosophical thought and of the prevailing psychology of philosophers. Earlier in this chapter, I contrasted the competing models of classical and current philosophical thought. The goal I have outlined has an appropriate application within only one of these approaches. It has a special attraction to those who endorse an age-old conception of philosophy, but who, at the same time, have become impatient with its more-than-patient, long-suffering, and unproductive accomplishments, and who are no longer willing to tolerate the continuation of a discipline that has produced little to nothing in the way of major reliable results to justify its existence.

Unfortunately, the psychology of human normality tends to thwart the achievement of this goal, not merely due to the existence of the rational bridge problem, but also because of another pervasive psychologically based human disposition that inevitably blocks constructive efforts.

### 1.2 What it takes to change a philosopher's mind

To put the point ... directly, there is no argument, valid or not, that can altogether silence a philosopher if he wishes to continue the discussion. However devastating the attack upon him, there is always something more he can say.... There is simply a reiteration of existing positions.... The wheels of argument spin idly, and the encounter reduces to a statuesque confrontation. Discussion degenerates into repetition relieved from time to time by name-calling.

– Henry W. Johnstone, Jr. (1964, pp. 481, 484-5)

The preferred beliefs of a great many philosophers, like the beliefs of the philosophically untrained majority, tend to be territorially fenced off: These prized beliefs are walled off by means of psychological denial, rejection, and distortion. In the normal course of life a philosopher's favored beliefs become, from his or her own point of view, intimately bound up with a strong—that is to say, often rigid—sense of personal, intellectual, and professional identity. Opposition to *self*-important and *self*-sustaining beliefs predictably is met with a great deal of resistance, which, as we shall see, is typically expressed through denial, rejection, and distortion.

The internal, reflex-arc process of defending one's fundamental philosophical beliefs is seldom self-conscious or self-consciously monitored, but

instead occurs as an automatic, habitual, self-protective process. For someone to question, and then perhaps even attempt to refute, the highly valued beliefs of many philosophers is, as we say, “to press that person’s buttons”: Such a philosopher is seldom able to imagine alternative and constructive ways of coping with and responding to what is experienced as a deeply felt assault on a system of beliefs which, very much like the human immune system, has been programmed to shield itself and to deflect potentially damaging criticism in order that it may survive intact and unscathed.

A philosopher who feels under attack will characteristically respond by attempting to deny that the attack is relevant to his or her system of belief, or by denying that the attack itself is justifiable, cogent, meaningful, logically consistent, or otherwise applicable. The overweening, single-minded purpose of a psychologically typical philosopher who is beleaguered by criticism is to reject the assault thanks to his or her tenaciously vested interest in defending the survival of that system of belief. To accomplish this, the assaulted philosopher will typically resort to distortion: He or she may systematically, and with sometimes considerable intellectual ingenuity, misconstrue, misstate, or misapply the attacker’s position, intention, use of language, or use of logic. This is defense by offense. But distortion also commonly takes the form of seeking surreptitiously to shift the ground of debate, of attempting subtly, and, if possible, imperceptibly by conceptual or linguistic sleight-of-hand, to change the formulation of the assailed beliefs, to “shape-shift,” to squirm out of tight spots—whatever distortions and contortions it takes to evade placing seriously in question, and potentially undermining, both the cherished system of beliefs and the philosopher’s own sense of self-identity.

This is an intensely serious and even grim business, and it is most assuredly *not* taken lightly, casually, or with humor by many philosophers. Unlike physicists and mathematicians, for example, who are compelled to revise their views—even though this may take time— by virtue of a shared methodology and strict standards of justification, the minds of many philosophers who are presented with rational argument and evidence are, as I have observed them, much like a duck’s back under falling water. This is not an accidental phenomenon, but one that pervades the discipline of philosophy and the psychology of its members. It is ubiquitously encountered at philosophy conventions, during presentations of papers, in interpersonal professional exchanges, and in philosophers’ responses to published work.

In such a setting, philosophical congresses seem often to be but caricatures of themselves. That one finds periodic points of similarity and even agreement, is small comfort when

compared with the magnitude of the gulf which separates chance islands of sound in an ocean of otherwise hostile silence. (Zaner, 1968, p. 61)

This phenomenon of *systemic, reflex-arc recalcitrance* serves to perpetuate and to reinforce the anti-result-orientation of philosophy. It is “systemic” because the philosophical beliefs and the philosopher’s commitment to them form a true system that creates the hard, impenetrable, reassuringly solid, cement foundation on which his or her personal, intellectual, and professional identity is built. The recalcitrance that is in view also has the property of a “reflex-arc” because it is the natural, automatic, defensive response of a system driven by survival interests. Philosophical belief system recalcitrance is no different on an abstract level than is the response of the immune system when it is challenged by a hostile pathogen: The immune system’s reactions (read: the defensive responses of the system of valued beliefs), and the reactions of the host organism (the philosopher) whose immune system (belief system) is potentially compromised, have a common dynamic. To study this philosophical-psychological phenomenon in the detail that is long overdue would take a book in itself. Here, only a few summary observations will be made.

The psychological characteristics that are common to members of any given discipline tend to fall into comparatively well-defined categories, which makes it possible to foresee through testing, with a certain measure of predictive reliability, how well an individual will fit and feel at home in a particular existing community of professionals. There are a number of such psychometric tests of interests and aptitudes, such as the Strong-Campbell Interest Inventory. The confirmed usefulness of such tests is an indication that distinct groups of professionals do tend—statistically—to share specific common characteristics. What characteristics are typically, that is, most commonly, shared by philosophers, considered as a group?

In other publications (Bartlett, 1986a, b), I described some of the most readily observed of these: In particular, they include a form of intellectual, and sometimes personal, narcissism, a variety of narcissism that expresses itself in a psychological-intellectual monadology: the ideological self-encapsulation of philosophers who, as Schilpp observed, “do not *want* to understand one another.” In a related context—religion—biologist Lewis Wolpert has used the apt phrase “*irrational persistence*” to describe the psychological predisposition involved in religious belief (Dawkins, 2006, p. 187). That so many departments of philosophy traditionally were combined with religion was not mere coincidence.

A profession whose members tend to share monadically, hermetically sealed systems of belief fences itself in within the boundaries of a set of internal limitations that obstruct constructive progress of the kind this study urges. Watzlawick, Beavin, and Jackson (1967) use the term '*imperviousness*' to mean what I have called '*recalcitrance*': The concern of these authors is the inability (and Schilpp adds the unwillingness) of an individual to make the other's frame of reference, position, or system of belief *real* in his or her awareness, and to think, feel, and act—if only vicariously with provisional sympathy—in ways that take the other's perspective genuinely into account. Such a person is, in short, impervious to what the other thinks and has to say. (It is the water-off-the-duck's-back phenomenon.)

When such a philosopher responds with irrational persistence, imperviousness, or recalcitrance, which he or she so often does in the face of potentially destructive objections, the resulting experience is one which Watzlawick, Beavin, and Jackson call '*disconfirmation of self by the other*' (1967, p. 91). This is again the phenomenon of undermined identity: "[W]hile rejection amounts to the message, 'You are wrong,' disconfirmation says in effect 'You do not exist' " (p. 86). This is of course hard medicine for anyone to swallow. To avoid it, denial, rejection, and distortion frequently come vigorously into play, as do assumed authoritativeness, professional pretentiousness, and polished skills in the rhetorical put-down: These together form the psychological ingredients of the "oil" that repels whatever water might fall on the duck's back.

Physicists and mathematicians, in contrast, generally and as a group tend to share decidedly different intellectual and emotional characteristics that enable them, and indeed eventually compel them, to "listen" to what their colleagues have to say. There are, of course, many potential bumps in the road along the way, since they—like philosophers and anyone carting along the baggage of a preferred belief system—do not easily budge from positions in which they have invested much of themselves and their lives. No one likes those who say "no" to his or her fundamental beliefs—until, of course, it becomes clear that saying "yes" to those beliefs was a big mistake. The standards of science and the criteria of mathematical rigor eventually manage to hold sway and in the end win out over the natural vested self-interests of individual scientists and mathematicians.

Unfortunately this isn't true of a great many philosophers. This observed fact is far from a chance occurrence, for a self-selection process goes on during anyone's education and professional training. A philosopher who "doesn't fit the current mold" will move to another discipline, unless he or she remains in a professional life that will fundamentally frustrate. Most fit well; those

who do not, experience what is, for them, justifiable frustration.

Given these psychological conditions, what, then, does it take—again, typically—to change a philosopher’s mind? At the very least it takes these things: (1) It requires an intellectual capacity conjoined by a psychological willingness both to consider a view not one’s own, and to measure against that view one’s own position through the use of agreed-upon standards of rationality and validity. (2) The first condition of course presupposes that there actually *exists* such a set of criteria that *is* agreed upon. Most importantly, (3) it is imperative if a philosopher’s mind is to be changed that he or she be one of the comparatively few individuals whose mental make-up enables that person to cross the bridge of rationality, and, in making that passage, form a persisting, non-provisional, non-tentative *internal commitment* to the rational results that are reached—even when these results overturn the philosopher’s own preferred beliefs: He or she must be *convinced* by reason on a fundamental emotional and behavioral level that strongly affects future thought and conduct.

It will already come as no surprise to the reader that, from the standpoint of the present work, a fourth condition needs also to be satisfied: For the purposes of philosophy advocated in this study, it is desirable that the mind of the philosopher that is to be changed be the sort of mind described in the left-hand column of the table included at the beginning of this chapter. That mind ideally should be the kind of mind that is both willing and wishes to discover universal, compelling truths. Without this overriding intellectual temperament and attitude, changes to a philosopher’s mind tend to be confined to delimited, often fragmented topics, limited in range of applicability, and falling short of the needs of a systematically integrated understanding.

Such an integrated comprehension at one time was the foremost objective of philosophical endeavor: to demonstrate the necessity of a conceptual worldview that spells out with clarity and justification how one is to perceive and think about the world, others, and oneself. The contemporary shift away from any such ambitious project considers that effort to constitute “systems building” and gives it a thumbs-down, preferring to embrace limited, pedestrian forms of analysis that are unable, and do not seek, to offer a total framework of understanding in terms of which one can think and live in self-conscious clarity. We shall return to the relationship between philosophy and a rational worldview in a later chapter.

Comparatively few professors of philosophy develop an interest and receive professional training in clinical psychology, and, as a result, serious studies by philosophers of the psychology of philosophers tend also to be few. Although I am aware that a number of the psychological conclusions de-

scribed in this chapter are unpleasant for many philosophers to hear, as they were for the author reluctantly to reach, we cannot afford to shelter ourselves from realities that obstruct meaningful progress by the discipline. The observations summarized in this chapter are the result of many decades of the author's interest, care, and concern for philosophy and its future. But the existence of this interest, care, and concern do not, I realize, take the potential sting out of the conclusions reached.

Philosophers—a very great many of them, as I have observed my colleagues and studied the works of our predecessors—generally lack the first of the characteristics identified above (the willingness genuinely to listen to and understand another's position); moreover, as Schilpp saw, they *do not want* to acquire and cultivate that disposition. As for meeting the second requirement (that philosophy reach an agreed-upon set of criteria and a unitary methodology which together determine the validity and acceptability of a philosophical position), the majority of philosophers today do not wish for this, either. Third, philosophers—and here they have much in common with everyman—comparatively seldom possess an intellectual-psychological constitution that enables, encourages, *and* compels them to cross the bridge of rationality and *commit* themselves to the conclusions which rationality demands. And, fourth, due to the swinging pendulum of fashion noted earlier, the great majority of contemporary philosophers do not possess the attitudes, values, and intellectual commitments that make for the kind of mind that is receptive to, and is determined to achieve, incrementally established, demonstrable, permanently enduring results.

In this light, we cannot avoid the conclusion that philosophy's long history of contention without end is far from being the result of chance: As long as the above underlying human conditions remain firmly in place among its practitioners, unceasing debates at cross-purposes to one another will predictably continue and philosophy will remain a discipline empty of substantive content. But, as I hope to show in subsequent chapters, this need not be so.

## 2

### **Putting Belief in Its Place: Its Psychology and a Needed Polemic**

[S]ome beliefs gain strength and become overly resistant to change. These are the beliefs that are central to our sense of identity, our cherished values, the social groups to which we align ourselves, or deeply held emotions. When scientific (or any other kind of information) challenges these beliefs, we are much less likely to change them because it would require a rethinking of our sense of self, what values we hold dear, the groups to which we have chosen to belong and what our gut feelings tell us.

– Geoffrey Munro (2011, p. 13)

**A**t least since Plato's Analogy of the Line—in other words, for most of western civilization's history of philosophy—we have been aware of the two extremes: mere belief and certain knowledge. Two millennia later, as one of the first philosophers to contribute in a major way to mathematics, Descartes was determined to go beyond belief in order to find a level of reliable knowledge. It is the same drive that motivates all science, all mathematics, indeed all disciplines that seek truth.

I do not include philosophy among these disciplines despite the fact that many individual philosophers have claimed to have sought the knowledge of what is true. Philosophy, as I have come to observe and evaluate the discipline, largely remains on the primitive level of belief as a result of the two main contributing factors we have encountered already: first, the dominant psychology of many philosophers themselves, and, second, the resulting majority displeasure over attempts to bring about disciplinary unity, to establish a shared methodology for the resolution of philosophical questions and problems, and to identify and then move beyond recognizable errors that have been made and dead ends that have been reached these past two thousand years.

The first of these factors—the psychology that dominates the profession,



which has been responsible for the floundering of philosophy, century after century, without reaching recognizably conclusive results—stands as a clear and so far inescapable obstacle to progress. For a psychologically focused and dispassionate analyst of the discipline, there is little doubt that the psychology of philosophy, the epiphenomenon of the individual psychology of philosophers themselves, has impeded constructive work. In studying the psychological core of the discipline and of its practitioners, we find a set of psychological traits that are constitutive of the philosophical “temper,” characteristics that have defined and continue to define the temperament and dispositions of many professional philosophers. Among these traits, as we saw in the last chapter, are intellectual and, frequently associated with it, personal narcissism, which lead to ideological self-encapsulation, belief system recalcitrance, and the propensities to engage in denial, rejection, and distortion when a philosopher’s views are challenged.

These dispositional traits result in three main consequences for the psychology of philosophers. These traits find expression in a governing, shared attitude among the greater number of philosophers, past and present, an attitude which (1) is resistant, as we’ve noted, to resolute attempts that would bring about disciplinary unification through adherence to a common methodology and the acquisition of demonstrable results that go beyond mere opinion; (2) is single-mindedly focused, through exposition, interpretation, and criticism, on what other philosophers have *said*, often to the exclusion of problem-centered, constructive, and incrementally progressing thought; and (3) is an expression itself of intellectual narcissism, an attitude that walls itself off as an area of human activity, limiting itself to internal discussion and disputation, and warding off attempts to develop philosophy beyond competing expressions and discussions of philosophically held *beliefs*.

...

*[M]an lives by those propositions whose validity is a function of his belief in them.*

– Gregory Bateson (Ruesch & Bateson,  
1968, p. 212; original italics)

Human belief takes many forms. Some beliefs are benign while others are nothing less than disgraceful, an insult to the human capacity to be intelligent,

to think rigorously, critically, with justification, and on the basis of evidence. The benign forms of belief are legion: Many serve us as habits do, such as belief that the sun will rise tomorrow, that the alarm clock will go off at the time for which it is set, that eating will satisfy hunger, etc. These are expectations we have been led to have on the basis of repeated experience, and our confidence or trust in them has accumulated in proportion to the extent that our habits and expectations have been formed and met.

But there are other varieties of belief that are held despite the fact that no factual or logically compelling evidence is available to support them; there are others that are held despite the fact that no evidence can *possibly* be given for them; and still others are adhered to despite the fact that to countenance the possibility of such evidence plunges us into paradox, self-contradiction, or absurdity. These latter three varieties of belief are what concern us here since, as we shall see in this work, many philosophical beliefs and positions built upon them exemplify those varieties. To different degrees they merit our intellectual disrespect, but more than this, they call for our resistance, impatient dismissal, or condemnation.

In his *Will to Believe*, William James mentioned beliefs that are believed because they are “in the line of your needs, for only by such belief is the need fulfilled” (James, 1979/1897, p. 34). In the pro-religious view he advanced, James considered such need-satisfying beliefs to be an expression of “wisdom” and “courage.” Here he was dead wrong. To trust in beliefs because in a variety of ways they are strongly gratifying can potentially transport us through the doors and into the asylum: They are, in extreme cases, the very sorts of beliefs which in our present society routinely classify people as mentally ill. At the far end of the clinical spectrum, delusional disorders are characterized by such beliefs: Those who invest their belief in the purely delusional can obtain a sense of self, purpose, and meaningfulness they rely upon. But even when not diagnosable as mentally dysfunctional, to invest our confidence in beliefs *because they make us feel better* brings us to the lowest depths of diminished rationality.

Both varieties of belief—those of the diagnosably delusional and those of the wishful thinker—are extremely resistant to change; their believers reject enlightenment when counter-evidence is offered; and when enlightenment *is* attempted, believers will often resist, sometimes fiercely and passionately, adhering all the more strongly to their cherished beliefs.

Beliefs that rest on no evidence, no possible evidence, or evidence that would, in principle, be self-undermining, are most frequently the kinds of beliefs that lead people to raise their hackles, stubbornly dig in to oppose any keenly felt challenge to their beliefs, and, in the extreme, can lead them to

become violent toward others and sometimes even toward themselves. To distinguish these particular varieties of belief with which I will be concerned from the many other varieties, I refer to them as *baseless beliefs*.

Epistemology is ideally an intellectually detached study of the limits and conditions of knowledge. It helps to train the willing and capable mind to discriminate among levels of claims to truth, and to assign to the different levels a range of degrees of confidence that we may justifiably invest in them. This is epistemology's more positive face, but it has another, as do the conjoined masks of tragedy and comedy, one representing Thalia, the muse of comedy, and the other, the grieving Melpomene, the muse of tragedy.

The more negative face of epistemology is judgmental and critical, increasing the sensitivity of our awareness concerning human claims to truth for which there is little, no, or no possible degree of confidence that we can have in them. This side of epistemology promotes discriminatory critical thinking. It is also the side of epistemology that potentially can, but most frequently does not, bring about deeply seated, lasting convictions in its students as they come to realize that much of what they have accepted on the basis of mere belief does not warrant the trust they have invested. They often will comment: "I can see that this result follows, but it's *slippery*." Their use of the word 'slippery' is significant; their use of the word is a kind of philosophical slip-of-the-tongue, for what is slippery is not the epistemological "result" they complain of, but it rather expresses the fact that their minds do not permit the result to gain traction in their own convictions (falling water and ducks' backs again).

This reaction is normal and is to be expected given the average philosophy student's, as well as the average professional philosopher's, intellectual and psychological constitution. There are no good *reasons* that can be offered for epistemology's failure, and the failure of its students, to form deeply rooted critical, discriminatory convictions which become part and parcel of their mental make-up—and which then *remain* after their epistemological training has come to an end. The failures here relate, not to reasons but to human *mental abilities*. The majority of students, as I have observed them from the standpoint of psychological assessment, confront what is for them the existence of an *unbridgeable bridge of rationality*. Like children whose minds have developed only to Piaget's level of concrete thought, they can attempt to engage in judgments, estimations, and calculations that presuppose a formal-operational level of thought; but they then see nothing amiss in providing a solution to a problem that, to a formal-operational mind, is absurd and ridiculous. The conclusions, to which their reason leads them, have, for them,

no persuasive reality and validity, and so their epistemologically unfounded, unjustifiable beliefs continue to survive with a life of their own.

### 2.1 Willful blindness

[W]illful blindness begins, not in conscious, deliberate choices to be blind, but in a skein of decisions that slowly but surely restrict our view. We don't sense our perspective closing in and most would prefer that it stay broad and rich.... [W]hat's most frightening about this process is that as we see less and less, we feel more comfort and greater certainty. We think we see more—even as the landscape shrinks.

– Margaret Heffernan (2011, pp. 20-21)

'Willful blindness' is a term that has been handed down to us from 19th century British law. It refers to knowledge or judgments an individual could have reached about a situation, or actions he or she should have undertaken, but despite what should have happened in light of such information, he or she nonetheless deliberately chose to remain ignorant and oblivious. Willful blindness is a widespread and familiar human trait, one that we should not be surprised to find also among many philosophers.

One of the identifying marks of willful blindness comes to the surface when a philosopher's beliefs are challenged; once this occurs, immediately and very often the tendency is for the philosopher to *increase* his or her advocacy of the cherished beliefs. As behavioral psychologists David Gal and Derek Rucker have discovered about people generally: "...individuals process information in a biased manner in an effort to maintain their beliefs. This outcome is particularly likely for beliefs that individuals view as important to their identity, such as deeply held political and personal beliefs" (Gal & Rucker, 2010, p. 1702)—and of course, one should add, deeply held philosophical beliefs.

This defensive attempt to protect preferred beliefs leads to a surprising result: The more evidence that is offered to undermine or disconfirm an individual's rigidly held beliefs, the more vigorously, stubbornly, and blindly will the individual be motivated to propound those beliefs (as shown, e.g., in the classical study by Festinger, Riecken, & Schachter (1956, 1957)). Moreover, given the contrariness of the human constitution, beliefs that are allowed to pass unchallenged tend ironically to be advocated *less* by their believers.

These observations apply directly to philosophical discourse: Attempts to challenge a philosopher's closely held beliefs will very often lead him or her

to increase the commitment he or she feels to those beliefs: The glue becomes ever stronger and more binding the more pressure is applied to pry a belief free from its believer. “Across experiments, individuals induced to feel doubt about their beliefs exerted more effort toward advocating their beliefs ... and expressed a greater likelihood to attempt to persuade other people of their beliefs” (Gal & Rucker, 2010, p. 1706). Expressed differently, the more a philosopher’s position is challenged, the greater lengths he is likely to go to in an effort to demonstrate that he’s right, willfully disregarding evidence and objections that might undermine his position. Related to the willful blindness found in law, the willful blindness of philosophers offers an additional source of understanding of the discipline’s pervasive psychology, whose consequence is the philosophical phenomenon of *systemic, reflex-arc recalcitrance* discussed earlier.

## 2.2 Giving belief its due—i.e., a bad name

[T]he higher one’s intelligence or education level, the less one is likely to be religious or hold “beliefs” of any kind.

– Richard Dawkins (2006, p. 103,  
quoting Paul Bell)

[B]eliefs are a luxury which we can get along without, and intellectual good form, as symbolized by Occam’s razor, demands that we should. For me, “belief” implies a metaphysical background, and metaphysical in the bad sense at that. It has always been a puzzle to me to understand why society treats a man’s beliefs with so much respect.

– P. W. Bridgman (1959, pp. 236-7)

One of the most discouraging things about the human bias in favor of belief is that many baseless beliefs are legally and socially protected, most especially the religious and political varieties: By means of the pressures of legal injunction and social taboo they are placed out-of-bounds to rational discussion and challenge. Most especially beliefs that are devoid of a speck of justification are judged to be “privileged”—which is to say that they are socially and politically granted a wholesale and indiscriminate *exemption* from rational scrutiny. All I need do in an argument is to state emphatically, “Well, *that* is my genuine *belief*”—(especially when it happens to be my *religious* or *politi-*

*cal* belief), and most discussions are brought to an abrupt halt in silent, respectful obeisance to the hallowed presence of belief that has been blessed with a privileged, blanket exemption from critical challenge. —For what else can those who would question such beliefs do? The questioner’s hands are tied. After all, “a person’s belief is a person’s belief,” and society accepts such belief as possessing an elevated, advantaged standing, immune to objection; “mere belief” is judged to stand firmly despite its all-too-shaky feet, which by social contract we agree to ignore. Respect is indiscriminately lavished upon such belief—uncritical respect which, if it were carried to its logical conclusion, would give the convictions of the mentally disordered the same “respectability” as those of any voiced belief for which evidence cannot be provided. This is the exemptive status accorded to *baseless beliefs*, a certain variety of which will form a later focus of concern in this book.

To offer undebatable respect for unfounded beliefs is to dignify them when they should instead be discarded and dismissed without hesitation or qualm. In any classification of levels of human cognition, baseless beliefs are relegated to the very bottom, among the arbitrary, capricious, unintelligent, uninformed, primitive superstitions and opinions expressed by the uneducated, the unreasoning, the bigoted, and that portion of the population rigidly committed to beliefs that have no basis other than the fact that they make their believers feel good.

But instead of relegating such beliefs to a lowly level of disrespectability, on the contrary and with incredulity we see that religion, in particular, but to a great extent also a significant number of philosophical views and opinions, have successfully promoted baseless beliefs to such an elevated level of respectability that to believe—fervently, uncritically, and indeed mindlessly—those beliefs that are most difficult to justify is perversely to exhibit virtue. Dawkins again has been on the right course:

Faith (belief without evidence) is a virtue. The more your beliefs defy the evidence, the more virtuous you are. Virtuoso believers who can manage to believe something really weird, unsupported and insupportable, in the teeth of evidence and reason, are especially rewarded. (Dawkins, 2006, p. 199)

We may be reminded of Giacomo Leopardi’s observation: “...there is nothing in the world so false or absurd that reasonable men will not hold true if the mind can find no way to believe the opposite and still stay at peace with itself” (Leopardi, 1981, p. 196)

Religious faith in the face of the strongest disconfirming evidence—e.g.,

the evidence of extreme human cruelty, disasters that cause indiscriminate and horrendous human and non-human suffering, illnesses and accidents that randomly strike down not only children, but those of great talent and achievement along with the best of people—is faith “to be admired.”

But the high esteem bestowed upon, and the injunction against questioning, baseless beliefs do not stop here. The privileging of baseless beliefs goes further: University of California at Berkeley psychologist Robert J. MacCoun summarizes one aspect of this phenomenon:

It’s the idea that whatever I believe, I believe it simply because it’s true.... This [psychological disposition involves] a “mental model”—a conceptual framework and mental representation about how something works that helps people make sense of the world. Once a mental model is in place, the mind tends to force new information to fit within it (quoted in Aschwanden, 2010, p. 33)

The privileging of unfounded beliefs exempts them from rational examination in these ways; they are respected with such uncritical and impassioned tolerance that they stand immune to serious challenge; they are placed out-of-bounds and out-of-reach by rational evaluation.

To give baseless beliefs the bad name they deserve is to recognize that they are the mental sustenance of the primitive mind, the mind at its lowest level of development and lowest ebb, which invests its trust in the phantasms that subsist in a delusional world beyond access by rational understanding, and whose resulting credulity in those fictions is elevated and praised as virtuous faith. Such beliefs have no place among intelligent, educated people, and certainly should find no niche in the minds of the specialists in that discipline whose members are reputed to have acquired the best of skills of rational, critical thought, careful analysis, and rigorous examination. I speak, of course, of philosophy.

If you, the reader, are to put such baseless beliefs in their proper, lowly place of inferiority, then it is necessary to cultivate a habit of *wincing mentally* whenever you hear someone, especially a philosopher, use the tell-tale phrase “I believe that...,” or its many variants: “I think that...,” “It seems to me that...,” “In my view...,” “It is clear that...,” “I suspect that...,” “I assume that...,” “It goes without saying that...,” etc., etc. To wince mentally whenever you hear such phrases is to become cognizant that you are in the presence of a game of *make-believe*—which phrase I use here in its several related senses: *manufacturing* a belief, *making* others believe it, all the while remaining in a

*made-to-order* universe based on nothing more than beliefs with no justifiable basis.

To wince mentally means further that you refuse to participate in the make-believe game, that you not only consider it a waste of time, but intellectually demeaning, a game that raises on a dais or places in the pulpit minds that are throwbacks to an archaic, undeveloped, pre-literate, pre-intelligent ancestry, before the evolution of rationality, before genuine knowledge was a human capability. If you recoil intellectually in this way, you become sensitive to conspicuous expressions of cognitive immaturity, and especially you then wince when faced by the disposition that insists upon the acceptability of a discipline which finds its substance in territorial disputes that arise from conflicting, baseless beliefs. This is what P. M. S. Hacker underscored earlier when he spoke of philosophy's two-and-a-half millennia of "irresponsible adolescence" (Hacker, 2009, p. 134), and perhaps what Timothy Williamson may have had in the back of his mind when he referred to "the intellectual immaturity that holds philosophy back" (Williamson, 2007, p. 8). Immaturity of this kind parallels the less developed level of mental functioning that becomes evident when comparing people who are, in Piaget's terms, concrete-as opposed to formal-operational.

To wince mentally in the above sense expresses a recognition that psychological recalcitrance to examine and dismiss baseless beliefs is itself an expression of a primitive level of cognitive development, indeed a sign of the mental savage for whom the occurrence of facts that conflict with strongly held beliefs is unacceptable and not to be tolerated without a fight, for whom new evidence is met with indignation and outrage when the evidence suggests that widely held beliefs should be thrown into the trash bin, for whom the agreement of other believers who concur with what is already believed reaffirms and buttresses the unfounded confidence that those beliefs are right and true, for whom results that are simply disliked are results that need to be explained away, and for whom, when presented with facts that are in conflict with their belief systems, a compulsion is experienced to scratch for any imaginable rationalizations that can be manufactured for holding such beliefs. We are reminded of Michael Polányi's closely related conception of *Weltanschauungen* that have a peculiar resilience and immunity to outside criticism; questions that fail to conform to the accepted and preferred conceptual scheme are rendered ineffective by the reflex judgment that they are undeserving of an answer (Polányi, 1952).

The word 'polemic', from the Greek '*polemikos*' meaning hostile and warlike, has acquired an unjustifiably bad connotation. The word has succumbed to popular distaste that is now enthusiastically directed against any



attempt to derogate the beliefs, opinions, and attitudes of others—all of whose mental states are, in a society fostering political equality, democratically elevated to a privileged status that confers equal merit. Here, in this chapter, I do not apologize for the present plain-spoken polemic against baseless belief—for if, as I am convinced, our baseless respect for, and indeed our baseless privileging of, unfounded belief has functioned as one of the major blocks to philosophical progress, and indeed as an obstacle to the progress of civilization itself, then both judicious hostility toward, and opposition against, unfounded beliefs are legitimately called for, and indeed are greatly overdue. More than any other human propensity, our species' passionate and willful commitment to preferred and baseless beliefs has been responsible for bringing about by far the greatest amount of destruction of life and happiness.

'*Polemikos*', akin to the Greek '*pelemizein*' meaning 'to shake', is precisely what philosophy (as well as the rest of humanity) has for centuries stood strongly in need of—a good, sound shaking sufficient to rouse even the most unconscious—long before Kant realized that Hume had shaken him awake from his “dogmatic slumber.” (Kant did not use the word 'shake', which may have been an oversight.)

In this brief, direct, and unapologetic polemic against the privileged role of unfounded belief in human society and in philosophy, in particular several psychologically based facts have been identified, which can be summarized as follows:

First, baseless beliefs very often serve general purposes important to the individual and the group: They bring satisfactions of various kinds, and in particular they help situate the person within his society and give him a mental ecological niche within which his or her personal, intellectual, and professional identity may reside comfortably. Second, baseless beliefs frequently serve people in their attempts to understand information about themselves, others, and the world. People make sense of these things by means of their systems of belief, which, as we've seen, are of a kind that is often very difficult or impossible to challenge and change. Third, we have seen how strongly held beliefs, within those who invest an inflexible commitment to them, acquire an immune system of their own, a systemic array of defenses that serve the self-preservation of those beliefs so that they may survive intact and unblemished. Most importantly, these defenses are able to achieve their purpose while simultaneously strengthening the blind trust, the reassuring confidence, that it is the psychological role of such beliefs to provide. Fourth, it does not matter should the beliefs be known to conflict with fact, they will be believed if they are what people *want* to believe. And last, direct challenge to entrenched baseless beliefs will frequently lead to willful blindness, outbreaks of

often violent emotion, and an increasingly fervent embrace of the challenged beliefs.

When encountered among professionals in the discipline of philosophy, these five human propensities should not be expected to be attenuated, although they may be transformed, or as Freud might say, sublimated to a higher level. The first propensity, the human satisfactions derived from a system of beliefs, is familiar to philosophers, who invest energy, determination, and often pride in their work in developing and elaborating a system of belief. Once a framework of belief has been established, a philosopher will apply his or her beliefs in filtering, or to use a more neutral expression, in “processing” the expressed thought of other philosophers. But in this processing, the systemic defenses of the philosopher’s system of belief typically make recourse to denial, rejection, and distortion in order to protect and preserve that system of belief. In this, the great majority of philosophers embrace the beliefs they prefer, and when pressed by challenges to those beliefs, many will make recourse to an array of rhetorical devices, such as undercutting the credibility of the opponent, discrediting the views expressed by the opponent’s position, and even sometimes seeking to place the opponent in a position of intellectual mockery, tactics that are intended to weaken the opponent’s public standing through humiliation or rhetorical put-downs, one-upmanship, professed authoritativeness, implicit or manifest ridicule, etc. These are a few of the characteristic forms assumed by outbreaks of undisguised strong emotion which take place in many exchanges among philosophers.

### **2.3 The rationale for such a polemic**

My purpose in describing the largely psychologically based predispositions of philosophers and to denigrate philosophy’s privileging of belief is not to express philosophically focused misanthropy. My purpose is rather twofold: to help to answer the question why, after more than two millennia, philosophy has so little to show for itself, and to urge, if we are to be professionals dedicated to the pursuit of truth, that we renounce baseless beliefs. Given that each of the disciplines attracts to its membership people who tend to have specific definable major personality traits, interests, and predispositions in common, we must recognize that those traits that typify the philosophical membership are vital to its capacity—or its inability—to evolve productively. Repeatedly in the philosophical literature we find philosophers raising the question why so little of philosophical consequence has come from the efforts of well-equipped intellects who have labored, often with great diligence and perseverance, during most of civilization’s history. To my knowledge, when such

questions are asked, answers are left in suspense, or else tentative, cautious, and nonspecific answers are suggested with gentle diplomacy. This state of affairs has not helped to move us along toward a more promising, constructive path.

It is of course possible that philosophy may deal with certain problems and questions which, by their very nature, are unsolvable and unanswerable in any conclusive way. But to establish (conclusively) that this is indeed the case would be tantamount to providing a Gödel-like proof of philosophical unprovability and incompleteness. Such a proof is not outside the limits of human imagination, although it would need to be of a meta-philosophical variety, standing above the level of routine philosophical discourse and thought, or else, paradoxically, such a proof could not be conclusive because it would become self-applicable and hence self-falsifying.

The speculative possibility of such a proof aside, we should recognize that the very psychological reality that this chapter has described, as long as that reality remains unchanged, creates obvious, serious, and perhaps inevitable and inescapable human obstacles that confront any attempt to reach philosophically conclusive results, even those concerning unavoidable philosophical inconclusiveness.

In the next section I turn from these preliminary reflections on the internal psychological limitations of philosophy to look at the one path philosophy has taken which is most closely related to the later technical content of this book.

## 3

### Turning Away from the Linguistic Turn: From Theory of Reference to Metalogic of Reference

#### 3.1 Through the looking glass: The superhighway of philosophy

Let us engage in a bit of allegorical fantasy. Think of the course of philosophy through its long history as a superhighway (an *elevated* and *high* way!), which meanders across a varied landscape that stretches in all directions. As we drive down the highway, here and there we pass exits that lead off the highway. Periodically we see signs informing us that we are on the “Highway to Universal Philosophical Truths.” But as we drive further, these signs appear less and less frequently.

There is a marker along the side of the highway that tells us we are entering the 20th century’s section of roadway, and then we can see many exits ahead lined up in closely packed succession, receding into the distance, sign after sign, exit after exit. We pass an exit marked “Existentialism,” another for “Phenomenology.” Among others clustered together in rapid succession, there are exits for “Pragmatism,” followed by “Positivism,” and then “Analysis” with an exit with multiple branches.

But it seems strange. A little farther on, we begin to see signs that look very much like the ones we passed miles and miles back, when we saw signs for “Metaphysics” and “Ontology.” But here we see exits with signs that say “Metaphysics, analysis of” and “Ontology, theory of.” There is even an exit with a sign that reads “Repaving project in progress” and beneath it, written in small letters, “Transcendental Philosophy, apparent exit.” We wonder for a moment what in the world an “apparent exit” could be—either it’s an exit, or it isn’t! But then being philosophers, we shrug our shoulders, think of other things, and continue on.

There are many more exits than I’ve listed, one after another, all marking turn-offs from the highway that was, so we thought, to take us to universal philosophical truths.

We continue driving with our characteristically indomitable philosophical patience, and cross what feels like an interminable expanse of flat, arid desert: featureless, monotonous, and quite boring. We pass many more exits: There is one marked “Linguistic Turn-off.” Along the side of the highway, there are what appear to be old roadside grave markers for “Phosphorus” and “Hesperus.” Each stands at the beginning of a separate exit, the first exit marked “The Morning Star,” and the second “The Evening Star,” but as we drive by we can see that the two exits converge into a single off-ramp. After we pass this pair of exits, we come upon another grave marker, for “The Present King of France,” closely followed by a smaller grave marker, for “Fido” and another for “Pegasus.” This is followed by yet another pair of exits, one marked “Scott” and the other “The Author of *Waverly*.” Like the earlier pair of exits for the morning and evening stars, these two exits also converge into one.

Soon after, we come upon a very large boulder on the side of the road into which has been chiseled in large letters, “Dr. Johnson’s stone.” There is a narrow exit to a small parking lot with a sign, “For those who wish to kick the rock.”

Some miles later another exit marks a turn-off to “Relativism.” And then we pass a series of off-ramps, all in dense succession, for “Multiculturalism,” “Hermeneutics,” “Structuralism” (not surprisingly, immediately followed by an exit, “Post-structuralism”), and then a sign in flaming Pink, written in flowing script, as we pass the turn-off for “Feminism, philosophies thereof.”

Ahead of us, and for the first time, the highway is jammed with traffic. The right lane is backed up, bumper-to-bumper for several miles. We take the left passing lane (this being an allegory primarily for American readers) and drive by the congestion easily. We soon see the reason for it: There is but a single exit marked with a sign that, unlike the other exit signs we’ve seen, lists several turn-offs that have been grouped together: “Epistemology,” “Philosophy of Mind,” and “Cognitive Neuroscience.” Cars attempting to take the exit have come to a complete standstill. Some of the drivers have gotten out of their cars and are clearly angry, faces red, gesticulating and shouting at the others. We’re happy not to be among them and drive on.

A little later, we pass a very odd exit, marked “Quadruplicity Drinks Procrastination.” We don’t know what to make of it, shake our heads, and someone in the car says it reminds him of an exit for “Zzyzx Road” which he saw when driving through the Mojave Desert.

The desert has become especially hot and we turn on the car’s air-conditioning.

As the miles roll by, we notice there are fewer and fewer exits, and almost no vehicles now on the road, which is almost empty.

We then pass under an overpass that bears a wide sign in large letters: “WARNING: BEYOND THIS POINT THERE ARE NO MORE EXITS.” The highway stretches ahead below the overpass and recedes in a straight line far into the distance.

We feel a bit worried and insecure, but continue on the highway, which, although it appeared straight as a standard meter stick as we drove under the overpass, now—and this is hard to describe—has begun to turn back on itself. It is something like driving on the surface of a Möbius strip.

The arid desert that was devoid of vegetation has, however, given way to lush clumps of forest interspersed with meadows in which many wildflowers are in bloom. In the middle of a large patch of wildflowers, we’re surprised once again to see, for the first time in very many miles, the once-familiar sign informing us that this is the Highway to Universal Philosophical Truths.

Here and there we also see very small diamond-shaped signs—and they really are quite small and hard to read even when we slow down—staked into the ground at regular intervals along the sides (or perhaps it has only one) of this recurving roadway. All of them say the same thing: “This is an invariant principle without which \_\_\_\_\_ isn’t possible.” The blank has either not yet been filled in, or the text has been worn away by winds that seem to be perennial here, or perhaps the letters in the blanks have faded so much as to be indecipherable.

And then we notice another strange thing: Just outside the line formed by the succession of the small signs, a fence of barbed wire stretches tautly parallel to them, mile after mile. Signs hang on the topmost wire that say “Trespassing prohibited. Violators face charges of meaninglessness.” Again, we’re not quite sure what to make of this, though some of us in the car are reasonably sure we know.

As we drive on, we gradually come to realize that we’ve seen the same small diamond-shaped signs repeatedly for a good long time now. There are no turn-offs, as we were cautioned some time ago, and then we come suddenly to realize in a philosophical epiphany that the highway has reached its end, but an end of a very peculiar sort, since the pavement continues as far as we can see.



### 3.2 An exit not taken

The pervasive and generally unquestioned assumption among philosophers of language today is that *concepts* are elusive, difficult to pin down and analyze, and hard to define, whereas language is concrete, specific, and more easily

made an object of public, collaborative (or contentious) study. Language, unlike concepts which it may express, is made up of letters, words, sentences, punctuation marks, and the rest, while it has a discernable grammar that can readily be studied. It lends itself to those who have a taste for the concrete, for material detail, who feel intellectually most comfortable on a less abstract and often less abstruse plane than the conceptual. Conceptually oriented analysis is theoretic, and not reducible to a study of words, phrases, sentences, and how these happen in fact to be used during any particular period—above all, by the ordinary speaker of the language. There is a political element, too, which should be obvious, for American egalitarianism resists the elitism of the intellectualism of highly theoretical thinking in philosophy, whereas it is at home in the democratic study of ordinary language as used by the down-to-earth, feet-on-the-ground regular person. Ordinary language and the uses it exemplifies have therefore been raised—much like the elevation of psychological normality by Western psychiatry to serve as a gold standard of good mental health (cf. Bartlett, 2011)—from the level of the merely ordinary, and have come to be celebrated as constituting a gold standard of philosophical method, and, as a result, these have been embraced for some decades by mainstream U.S. and British philosophy as the most satisfying of possible approaches to philosophical questions.

Philosophy of language has attempted to sound the death knell for a large and varied set of objects: ideas or concepts, abstract propositions (as opposed to concrete sentences and utterances), meanings, and even abstract knowledge. In their place, those who take the linguistic turn content themselves with a Quinean taste for desert landscapes—in other words, with words and more words, which do not signify in any important, potentially “mystifying” and higher sense, but are the straight-forward, analyzable expression of speech habits, of having learned how to use words, names, and phrases, and how to compile them into grammatically acceptable sentences, some of which are regarded as having “values” like truth, falsity, and even meaninglessness. (Many language analysts, however, have generally given up on the “*concept* of meaning” and have moved on to other topics—so that a state of being “devoid of meaning” makes use of a word they no longer consider especially useful.)

The desert of linguistic analysis not only possesses a sterility, in which some philosophers may appreciate the aesthetically attractiveness of surgical sanitation, it can be and often is, for some philosophers, monotonous, pedantic, and boring. But “boring” and “pedantic,” like “sterility,” can, also for some, have an unapologetically positive side: As one contemporary language analyst wrote, without any sign of self-conscious embarrassment, “The fear of

boring oneself or one's readers is a great enemy of truth. Pedantry is a fault on the right side" (Williamson, 2007, p. 288). If boring pedantry is what graduate students of philosophy have to look forward to, in my view we face a sad situation indeed.

This book turns in the opposite direction and once more legitimates and breathes life back into those abandoned thoughts, ideas, propositions, and meanings. Readers tired to the point of mental exhaustion, exasperation, or intellectual nausea when they encounter yet more discussions of "Hesperus," "Phosphorus," and how to understand whether the present King of France is or is not bald, will be comforted to know that they'll read about none of these in future chapters. In the process, language is demoted to the merely contingent, to the incidental instrumental means that we often resort to in expressing thoughts, ideas, propositions, and meanings, recognizing that some of these are expressed more effectively and eloquently in individual affective response: in art, music, poetry, and other arts, and in many other ways, including silence.

In this book I recognize and do not object to the naive view regarding the nature of language and meaning: that language is used to refer to objects in the world; that it is used to express thoughts, emotions, imagery; and that it acquires meaning through consensus, habitual use in the context of public goals, intentions, etc. But I stop there and see no convincing reason—or evidence—to accept that the analysis of a contingent, happenstance tool of expression—changeable, natural, disorderly, equivocal, vague, often inconsistent ordinary language—offers a most promising method by means of which to obtain knowledge of the functioning of our minds or of the nature of reality.

Philosophy of language, in its attempt to pin down, analyze, and systematize the natural resilience, multi-purposefulness, and ambiguities of ordinary language, has often ended up perverting the above-mentioned naive conceptions of language through oversimplification, reductionist thinking, myopia, and philosophy's love for technical distinction-making and the development of an unnecessarily elaborate terminology accessible only to specialists. In the author's estimation, the increasingly detailed, complex, and convoluted controversies among philosophers of language stretches a naturally evolved means of communication on a Procrustean bed of theoretically contrived constructs, and in this process the shape-shifting, adaptive, elastic, inherently extendable, malleable, unruly nature of natural language slips through the sieve of analysis. This should come as no surprise. In the process of such stretching and technical crystallizing, philosophers of language closely attend to what their fellow language analysts have said and written, and *their* words and



phrases then become *the* subject-matter, and not the nature of thought, experience, or reality.

In the framework developed here, which is open to a study of concepts and meanings, there is no static “entity” that underlies and is to be equated with linguistic meaning. Natural language can be molded to a great many expressive purposes that we may wish to communicate, so it is at once simplistic, and, if we can learn one lesson from the experience of philosophy of language, it is also misguided to wish for a well-defined circumscription of the nature of the many layers and kinds of meanings that language is capable of expressing. Natural language is not reducible to any one thing or any set of enumerable things, for there is no limit to the rich and evolving vocabulary and number of sentences such a language can be used to express.

To understand and perhaps, for some readers, to sympathize with the justification here to turn away from language analysis, it can be useful to remind ourselves of the original historical and the then-persuasive motivation for making the linguistic turn to begin with.

The *belief* that natural language and human thought are intimately bound together has a long past. From Plato to St. Augustine, through the Middle Ages, until Kant, the trend among most philosophers was to uphold an instrumental view of language, as a *mere tool* that expresses or refers to concepts and facilitates our understanding of experience.

But by the time Kant was born in 1724, a quite different revisionary basis for philosophy of language had already begun to take root in France, where the study of language was elevated considerably by the so-called ‘*grammairiens-philosophes*’, who included Kant’s contemporary, César Chesneau Dumarsais (1676–1756), Étienne Bonnot de Condillac (1715–1780), and their successors, the *idéologues*, whose shared passion was to formulate what they called a ‘universal grammar’ that ideally would express what is common to all individual languages.

Thanks to the work of the *grammairiens-philosophes* followed by that of the *idéologues* the study of language became an end in itself: Condillac *believed* (wince here if you can) that the structure of language mirrors the structure of thought: He and his successors were so taken with this *belief* that the *écoles centrales* of France replaced the classical chairs of logic and metaphysics with chairs devoted to the study of universal grammar. Although these radical academic changes in France were only to last until the end of the 1700s, the *belief*—a belief which had no basis in evidence, that a study of language promised important knowledge about human thought and even physical reality—has persisted, but not in France. The French eventually came to their senses, and language analysis migrated to Britain and the United

States, while continental philosophy headed in other directions. It may come as a surprise to many readers that France was the originator of language analysis, so closely has philosophy of language come to be associated with Anglo-American philosophy.

Foreshadowings of contemporary ordinary language analysis appeared early. Johann Gottfried Herder (1744–1803) *proclaimed* that language *determines* thought. Wilhelm von Humboldt (1767–1835) *believed* he had “discovered the art of using language as a vehicle by which to explore the heights, the depths, and the diversity of the whole world” (letter to Wolfe, 1805). And J. S. Mill (1806–1873) described metaphysics in his book, *A System of Logic*, as “that fertile field of delusion propagated by language.”

The belief that language (perhaps) determines thought was later picked up and developed in the 20th century by Edward Sapir (1884–1939) and Benjamin Lee Whorf (1897–1941), and then by their successors, who called this belief the ‘linguistic relativity hypothesis’. Its strong version claims that a natural language determines what users of that language *can* think and experience, while its weak version claims only a relation of influence of language on thought and experience. The empirical jury is still out whether either belief can be substantiated to any significant extent; some studies appear to show a weak degree of influence of one’s native language upon, for example, one’s color discrimination skills, but other studies do not confirm this, and some suggest that such influence can be overcome by appropriate enrichments or extensions of the language in question. (For references to the literature, see Hardin & Maffi (1997) and Pitchford & Biggam (2006).)

As the cement of philosophical fashion to see in language analysis a reliable method to investigate thought, experience, and reality set and solidified into a monolithic block—with the paucity of supporting evidence shrugged off—few voices dissented. One dissenter was anthropologist and linguist Franz Boas (1858–1942) who held a Ph.D. in physics. In his *view*, any natural language is capable of expressing the same content, but often using varied means. He recognized the inherent, adaptive, flexibility of language and the endless ways it can be extended to meet new conditions. Noam Chomsky (b. 1928) in a parallel way advocated the *belief* (once again) that language does not affect thought processes that are to be found among all people. Belief in linguistic relativity—among anthropologists and linguists—has since then generally fallen out of favor among linguists and anthropologists, but not among many philosophical language analysts.

The still unfounded belief concerning an intimate relationship of language to the structure and nature of thought, experience, and the rest of the world has survived in philosophy, and even in computer science where the simple-

minded belief has been advocated that anyone who learns a particular programming language will come to “think” in that language in a manner that governs how they develop computer programs. (For example, the work of Kenneth E. Iverson and Paul Graham.) This is no more, no less, profound an observation than the truism that learning to count in the decimal system will lead one to count in the decimal system (while counting using a base other than ten is easy to learn and ubiquitous in computer science). Nonetheless, here the strong version of the linguistic relativity hypothesis again raises its banner.

When we look closely at the past century of work in philosophical language analysis, we continue to find an abundance of belief statements, but virtually no evidence for a great many of them. Frege maintained the belief (for philosophical claims have nearly always rested on belief) “[t]hat a thought of which we are conscious is connected in our mind with some sentence or other is for us men necessary” (Frege, 1979, p. 269). (Why this is “necessary,” he never explained.) Some philosophers express the belief that thought is internalized speech, so that to understand thought, we should or must investigate speech. For example, “philosophers can contribute more by investigating *discourse* about mental states than by investigating the mental states themselves.... [A]ttention is focused upon plausible possible speech situations in which one person talks to another about the mind of a third” (Woodfield, 1982, p. ix). Another pro-language philosopher advocates the following belief:

[H]owever imperfect it may be to have to work with an intermediary, some linguistic vehicle is *required* to go proxy for the Platonistically conceived realm of thought; and the study of the structure of thought *had* to proceed, with due caution, *via* the study of the semantic structure of language, itself something which was not fully present to the naked or untutored eye.... [A]ny account of thought *must* go *via* an account of the language in which that thought is expressed. (Smith, 1995, pp. 27-28; italics added, except for the two ‘*via*’s).

—Why is a “linguistic vehicle” *required*? Why does the study of thought *have* to involve the study of semantics? Why does cognitive content and a medium for its expression *have to be* essentially related? Surely we must question whether it is necessary to know anything about the semantic structure of a particular word or sentence in order to attend to an abstract proposition that

can be expressed in multiple languages, and may even be expressed non-verbally (think of the very precise instructions communicated non-verbally by a conductor like Toscanini). Again, the asserted beliefs that one *must* study language in order to understand thought, experience, and reality are left to float about in a capricious mental space, entirely unfounded. The last-quoted philosopher concluded: “For creatures like us, reference is brought about and sustained by linguistic cognition” (Smith, 1995, p. 36). Again, why is or must this be so? He doesn’t say.

An analogy forces itself into our discussion at this point: Let us suppose that we come across a toolbox containing a variety of tools. What credibility would we give to the claim that by analyzing the tools we can infer or otherwise come to know the structure of the objects (machines, homes, bridges, sculpture) that the tools are, or can be, used to make? The contents of a toolbox are so very convenient to handle and study: They are concrete, down-to-earth, and specific in their construction. They can be made into objects of detailed study. But it takes little mental acumen to recognize that there is an exceedingly wide gulf between a study of instruments for making things, and the nature of the things that can be made with them. The old metaphor of the drunk searching for his keys in the light cast by a streetlight, simply because there is more light there, is apropos.

One philosopher of recent vintage protests against such a loss of research respect in studying the contents of the box of tools. He asserts the following belief:

Philosophers who refuse to bother about semantics, on the grounds that they want to study the non-linguistic world, not our talk about that world, resemble scientists who refuse to bother about the theory of their instruments, on the grounds that they want to study the world, not our observations of it. (Williamson, 2007, pp. 284-5)

But in urging the need to study our linguistic instruments of expression, this believer in the central role of language ravel together “instruments,” “theory of instruments,” and “observations,” each of which is decidedly different: The choice, for example, of a meter or yardstick is irrelevant to the theory of measurement in physics, while observation is not a simple concept, but has required an elaborate theory of its own. As I understand it, the passage quoted from Williamson seeks to defend his fundamental belief that a study of our “instruments”—that this, paying close attention to the semantics of language—ought, in some way, to serve at least as a partial arbiter of what is

true. As we shall have occasion to disentangle later on, this is far from true.

Philosophy of language has become obsessive-compulsive in its insistence that we focus on linguistic structure if we are to make any headway in understanding the subjects that philosophy has set itself to understand, or if we are to find our way out of the alleged fly-bottle that traps philosophers who have been bewitched by their misleading and mistaken uses of language.

In promoting this fashion that began in 18th century France to look to language as a source of philosophical knowledge, Quine has acquired the authoritative reputation as *Imprimatur* in his advocacy of what revealingly (in a linguistic slip of the tongue?) he has called a '*taste*' for desert landscapes, for *taste*—that is, *fashion*—is what is at stake. His *belief* is quoted far and wide: "the way to clarify our talk of ideas is not to say what ideas are, but to show how to paraphrase talk of ideas into talk about language" (Quine, 1987, p. 88).

And so, in book after book of philosophical language analysis, we do in fact now find such paraphrases as Quine wanted, often a great many of them. Their object, in Quine's words, is to make sure that our "feet are on the ground" (p. 88). In another publication, he elaborated upon the *belief* that he propounded:

It was emphasized by rationalists and empiricists alike that inquiry should begin with clear ideas. I agree about the clarity, but I balk at ideas. The British empiricists themselves balked at abstract ideas. *Nil in mente*, they declared in their orotund British measures, *quod non prius in sense*. They echoed their nominalist ancestors, for whom abstract ideas were *flatus vocis*—words, words, words....

Let us therefore recognize that the whole idea idea, abstract and concrete, is a frail reed indeed. We must seek a firm footing rather in words.... [John Horne] Tooke held that Locke's essay could be much improved by substituting the word 'word' everywhere for the word 'idea'. What is thereby gained in firmness is attended by no appreciable loss in scope, since ideas without words would have come to little in any event. We think mostly in words, and we report our thoughts wholly in words. (Quine, 1978/1977, p. 155)

And yet, even Quine finds it hard to dispense with the world (perhaps he would say only "the vocabulary") of the conceptual: In the same essay he makes such claims as "A body is *conceived* as retaining its identity over time between appearances" (p. 159); "[b]odies are basic to our *way of thought*, as

objects go” (p. 159); and “[i]f a man were twice in the same physical state, then ... he would *believe* the same things both times, and would have the same *thoughts*....” (p. 162). In the remainder of his essay, he goes on to use the expressions ‘mental life’ (p. 162), ‘mental events’ (p. 163), ‘motive’ (p. 164), ‘attitude’ (p. 165), ‘thought experiment’ (p. 167), and ‘conceptual apparatus’ (p. 168). —For shame! His *language* is filled with *conceptual-psychological* terms. They have not been paraphrased-out to reduce them to the ground level upon which he urges us that our feet must solidly be planted.

This alleged *need* to paraphrase ideas, concepts, or any conceptually based content has never been justified. It is, as Quine admits, a matter of *taste*: He is completely honest about this, admitting that “abstract ontology” is “far from *congenial* to nominalist *tastes*” (Quine, 1987, p. 299; my emphasis). And it is purely a matter of *belief* that the study of the contingent linguistic means offered by often primitive, changeable, and evolving ordinary language should provide us with knowledge of the structure of thought, experience, and perhaps even of the physical world. Here there is supposedly a bridge of a kind very different from the rational and moral bridges described earlier: It is a bridge from the structure of language to that of thought and that of the real world, a bridge that is purely imagined; it is based upon belief which itself is baseless.

Not only this, but here are several intermixed confusions and clear errors in the passages (and the thoughts they express) which I’ve quoted from Quine: Let us suppose that a psychotherapist seeks to treat the human propensity to think projectively (in the psychiatric sense: to assign to another what one wishes to deny in oneself). There would be a very considerable “loss of relevance” in substituting words for ideas, for such a therapy certainly does *not* have as its goal to persuade people who engage in psychological projection *to use language differently*.

Quine admits that not all thought is in terms of words (“[w]e think mostly in words”), and surely not all thought is verbally based: Einstein reportedly often thought visually, in terms of imagery,<sup>9</sup> so do I some of the time, so do

---

<sup>9</sup> In a letter to mathematician Jacques Hadamard, Einstein wrote: “The words or the language, as they are written or spoken, do not seem to play any role in my mechanism of thought. The psychical entities which seem to serve as elements in thought are certain signs and more less clear images which can be ‘voluntarily’ reproduced and combined.... [T]aken from a psychological viewpoint, this combinatory play seems to be the essential feature in productive thought—before there is any connection with logical construction in words or other kinds of signs which can be communicated to others.... The above mentioned elements are, in my case, of visual and some of muscular type. Conventional words or other signs have to be sought for laboriously only in a secondary stage, when the mentioned associative play is sufficiently established and can be reproduced at will” (Hadamard, 1954/1945, pp. 142-143).

many people. Most assuredly musicians and artists do not think in language much of the time. But, as we've seen, Quine goes further when he asserts his bare-bones belief that "we report our thoughts wholly in words." Has he never listened to the music of Bach, Beethoven, and Brahms? or been to the Louvre, the Sistine Chapel, or the Lascaux Caves? We find no words in the thoughts that were expressed by any of these. Many people and many human creations do not report thoughts in words—in fact, some seldom use words. Quine's claim that "we must seek a firm footing in words" belies not only a taste, but a *prejudice*, it expresses a claim that has never been justified, by him or by any other language analyst. It harkens back to the strong interpretation of linguistic relativity, which has never been demonstrated.

To conclude this section in which I have pointed a finger of shame at philosophy of language, it might interest readers to recognize the observations made more than a century ago by Fritz Mauthner (1849–1923), from whom the much younger Wittgenstein (1889–1951) borrowed rather too freely—some might go so far as to say virtually plagiarized—without giving credit where credit was due.<sup>10</sup> The following are some of Mauthner's claims, which most philosophers of language today believe, very mistakenly, originated with Wittgenstein:

Language, in Mauthner's view, is nothing more than the embodiment of certain rules of a game [*Spiegelregeln*] (Mauthner, 1901-1902, vol. I, p. 25). "Language is ... no object at all, it is nothing but its use" (vol. I, p. 24). "I must do away with language ... step by step—I must break each rung of the ladder as I tread on it" (vol. I, pp. 1-2). (Wittgenstein's had no qualms in taking over precisely the same, now famous, metaphor under his own name.) Mauthner claimed, before Wittgenstein hit his teens, that "the philosophers [sic] task is to ... free us from the spell of language. This will be the self-criticism of philosophy through the criticism of language" (Weiler, 1958, p. 85). In Mauthner's view, philosophy should engage in a critique of language, and

---

<sup>10</sup> Had Wittgenstein been intellectually responsible by providing due credit to Mauthner's ideas that he incorporated in his *Tractatus* and *Philosophical Investigations*, Mauthner's contributions rather than Wittgenstein's would displace many and perhaps most of the laurels that have been bestowed on Wittgenstein's head. Instead, Anglo-American philosophers have acquired the mistaken belief that they owe a great deal too much to Wittgenstein, and many are not even familiar with Mauthner's name.

No acknowledgment whatsoever of Wittgenstein's debts to Mauthner appears in the most often used dual-language, English/German editions of the two most often studied works by Wittgenstein (the *Tractatus Logico-philosophicus* translated by D. F. Pears and B. F. McGuinness, and *Philosophical Investigations*, translated by G. E. M. Anscombe). Wittgenstein makes only a single dismissive mention of Mauthner in his *Tractatus*, and that is to *distance* himself from Mauthner by saying: "All philosophy is a 'critique of language' (though not in Mauthner's sense)" (§ 4.0031).

by so doing the results would be therapeutic for philosophy (again, Wittgenstein's copied claim). For Mauthner, the end of this process is silence, contemplative quiet (again, imitated by Wittgenstein.)

Mauthner has been very considerably overlooked and neglected in the history of language analysis. Perhaps this is attributable to the fact that his major three-volume work *Beiträge zu einer Kritik der Sprache* [*Contributions to a Critique of Language*], published more than a century ago (in 1901-1902) has *still* not been translated into English.<sup>11</sup> No doubt there are reasons for this shortcoming—we shall speculate in a moment—but for whatever the reasons, Mauthner has been given short-shrift and has been virtually ignored by Anglo-American philosophers. (A few exceptions include Weiler, 1958 and 1970; and Bredeck, 1992.) If posterity is to be fair to Mauthner, though such fairness would come more than a century too late, the extent of Wittgenstein's unacknowledged borrowing from him needs to be brought into the light of day.

As for the possible reasons for his neglect, Mauthner perhaps did not invest the same degree of uncritical, blanket trust in ordinary language as did Wittgenstein. One wonders if this “departure from ordinary language faith” may in part have been responsible for the objectionable negligence of his work by Anglo-American philosophers. In Mauthner's reasoned view, any attempt to acquire knowledge about reality from the structure of language reflects nothing more than “word-superstition” [*Scheinwert*]<sup>12</sup>—in other words, what I have underscored as unfounded belief. “The man who claims to know the world through language is like the spider, who imagines he knows the palace in one of the concerns of which he has established his web” (Weiler, 1958, p. 81). “[The] critique of language will ultimately yield no knowledge of the world” (Mauthner, 1901-1902, I, p. 689).

Other possible reasons for Mauthner's neglect include that he was a Jew; that he lived in Berlin rather than then-philosophically-magnetic Vienna; that he did not confine himself to philosophy, but—heretical for a dedicated *professional* philosopher—had a sufficiently broad mind also to write journalistic and literary works; and, probably more than any other reason, that Russell loaned his considerable influence in fostering among his contemporary philosophers a blindness of credulity in Wittgenstein's authoritativeness, which

---

<sup>11</sup> Wikipedia recently added a comparatively brief article about Mauthner ([https://en.wikipedia.org/wiki/Fritz\\_Mauthner](https://en.wikipedia.org/wiki/Fritz_Mauthner)). Interested readers who cannot read the more extensive German Wikipedia article about Mauthner can consult Google's rough English translation through the Google Translate: <https://translate.google.com/>; then enter the following web address in the translation box, [https://de.wikipedia.org/wiki/Fritz\\_Mauthner](https://de.wikipedia.org/wiki/Fritz_Mauthner), and click Translate (accessed August, 2020).



conferred authority appears to have given credibility to Wittgenstein's conveniently quick dismissal of Mauthner's work in his *Tractatus*.<sup>12</sup>

In the foregoing discussion, three ironies stand out: (1) The genealogy of Anglo-American language analysis goes back to the now-archaic French enthusiasm for this way of doing philosophy, which the French fairly quickly decided was a dead-end, bypassing the philosophical highway's "Linguistic Turn-off." (2) Fritz Mauthner's major role as the source for Wittgenstein's key ideas has been generally unacknowledged, ignoring the fact that Mauthner originated a group of theses about language that would later become Wittgenstein's unearned hallmarks, while Wittgenstein's absorption and copying of Mauthner's way of thinking has been overlooked and forgotten by nearly everyone. (3) Mauthner recognized that the bridge between the structure and nature of language, on the one hand, and human thought, experience, and reality, on the other, which language analysts *believe in*, is nothing more than "superstition." The indirect evidence supporting this claim by Mauthner lies, of course, in the absence of evidence that has ever been provided for this fundamental belief shared by many philosophers of language.

At this point, we shall take a different route, not only passing by the turn-off of philosophy of language, but, more importantly in this study, also passing by its approach to theory of reference.

### 3.3 The overlooked variety of forms of reference

The topic of reference attracted a great deal of attention during the 20th century. In its most general formulation, reference has been conceptualized as a relation between language or thought and the world, and as a means by which objects are referred to. More comprehensively, reference has been understood as a function of four relations among (1) a speaker or thinker, (2) a word or thought, (3) an object, and (4) an audience. The subject of reference and the study of theory of reference to which it has led have been almost exclusively confined to language analysis which has come to regard theory of reference as its territorial possession. Language analysts have studied the subject of reference in terms of various kinds of relations that may be identified between symbols and other symbols, or between symbols and things other than symbols. These symbol-based relationships have been discussed and analyzed under such familiar headings as meaning and denotation and their many cousins that have been distinguished during the last century's history of semantics. The historically familiar approach to the subject began with Frege

---

<sup>12</sup> See note 10.

and Russell, and was subsequently critiqued, amended, elaborated, and diversified by Strawson, Searle, Quine, Kripke, Putnam, Kaplan, Donnellan, and others.

Three competing basic theories (again read: beliefs) about the nature of linguistic reference have established the foundation for most of the variants that have evolved since. The three basic theories are Frege's theory of sense and reference, Russell's theory of descriptions, and Strawson's theory of referring. On this foundation several further views about the relation of language to objects have developed: Kripke's and Putnam's non-descriptivism, which in turn led some philosophers of language to develop a causal theory of reference (according to which a causal chain connects a term with an object); Kaplan's view of direct reference; and the minimalist view of reference proposed by Rorty, Brandom, Horwich, and others. These competing views (again: beliefs) describe possible ways of construing linguistic reference. It will not be relevant to the subject-matter of this book to explain and discuss these views; there is already an abundant literature about the views of these philosophers, which will not concern us since—to re-emphasize this—language itself is here relegated to its role as a mere accidental tool of expression, while reference as we shall understand it is a good deal more than linguistic reference. Reference takes many forms, and its multiplicity and variety should help to liberate us from the narrow constraint that gives priority to the study of language.

There are, in fact, so many types and kinds of reference that it would take a book to attempt to inventory them. However, if there is any consensus about reference it is the claim that to think or talk about anything is to refer to it. — Take any discipline, and you will find indigenous forms of reference which are employed there. Consider anything that can be the subject of attention, and reference plays a role. Reference may be made to facts, abstract objects, people, their cognitive and emotional experience and its behavioral expression; to characters, situations, and stories imagined in fiction; to processes of reasoning, stream-of-consciousness monkeys of the mind that come and go at will, and the content of dreams and the feelings they arouse in us; to ideas of time, space, nondenumerable orders of infinity; to anything identified, recognized, or remembered; and, to bring a potentially interminable listing down to a very ordinary level that whets Quine's taste, there is the variety of reference found in Social Security numbers, which embody a system of encoding that refers to the time and place an individual obtained his or her number. At the same time, a Social Security number ideally, and most of the time in reality (short of identity theft), serves uniquely to identify an individual person.

In addition to these more straightforward forms of reference, there is self-

reference, or reflexivity, which is a subject unto itself and of particular importance in the present work. There are many forms of self-reference, of which I will list a representative set with some occasional examples. They include the linguistic reflexivity of indexical signs, egocentric particulars, and token-reflexive words; there is semantical self-reference; tautological self-reference (as an example, the reflexive incorrigibility of certain claims about immediate experience); set-theoretical reflexivity (which led to an array of paradoxes: the familiar Burali-Forti paradox, Cantor's paradox, Russell's paradox, the Richard paradox, the Zermelo-König paradox, Berry's paradox, Grelling's paradox, etc.); pragmatical, or performative, self-reference (studied by John Passmore, Henry W. Johnstone, Jr., J. L. Mackie, and others); meta-logical self-reference (about which we will learn more later); reflexivity in artificial intelligence (self-correcting systems, self-regulating systems, self-organizing systems, systems capable of self-initiated learning, self-reproducing systems, etc.); reflexivity in physical theory (in particular in quantum mechanics and general relativity); reflexivity in topology (lines, surfaces, and volumes that recur upon themselves); biological reflexivity (self-replication and self-organization); self-reference in political science (e.g., self-limitation of political power); reflexivity in law (laws and policies that permit self-amendment; self-referring laws, etc.); sociological reflexivity (e.g., the problem of reflexive prediction); self-reference in economics (reflexive monetary adjustment theory, the dynamics of self-fueling inflationary and deflationary systems, etc.); reflexivity in game theory and decision theory (rules permitting self-modification, decision methods relating to the ordering of preferences); reflexivity in anthropology (perhaps the most famous variety here is the linguistic relativity hypothesis which we have already come across, a hypothesis about language which itself is reflexively expressed in language); self-reference in mythology and theology (self-embodiment of a deity in a universe created by it, the sometimes purported reflexivity of the predicate of perfection in ontological arguments); self-reference in literature (self-begetting, self-describing, or self-reflexive works); in science fiction (closed loops in time, paradoxes of self-identity, etc.); in music (reflexive loops in such musical forms as the fugue and canon); reflexivity in art (self-depicting art, as in a scene containing an easel and canvas, on which the same scene recurs to the limits of resolution); reflexivity in fractal theory (the elaboration of self-evolving structures as the result of a formalized computational rule whose output is successively and repeatedly made its input in a never-ending cycle); in psychiatry and psychology (reflexivity gone awry in many forms of alleged mental illness); self-reference in information theory and general systems theory (the familiar phenomena of feedback and feedforward); self-reference in

hermeneutics, the theory of paradigm change, and the theory of research programs; and then there is reflexivity in neurophysiology (e.g., self-referential thoughts and feelings such as pain and anxiety, feedback and feedforward loops in the brain that may permit self-conscious awareness, etc.). Finally, we cannot forebear mentioning the species of self-reference habitually found in academic scholarship: the use of citations to one's own publications, which are self-references of a specific, concrete sort.

I will not attempt to define these multiple forms of self-reference or reflexivity here, which would take us far afield.<sup>13</sup> The reason they are worth highlighting is that relatively few philosophers, even those who study theory of reference through language analysis, recognize in their writing the extremely great variety of forms of reference and self-reference.

With this partial list of varieties of reference before us, one further major class must be mentioned, since it seldom is given due attention, and that is non-linguistic reference, reference that occurs non-verbally. We do not often confront and acknowledge the expressive limitations of language, or the fact that a great deal that must be subsumed under the category of "reference" occurs on a non-verbal level. This is an empirical fact: language is not coextensive with the domain of reference.<sup>14</sup>

Psychologists of communication Bateson and Jackson used the terms 'digital' and 'analog' to distinguish verbal and non-verbal expression:

[W]hen we consider the plight of man, we observe at once that he has great paucity on the digital side. There is probably no systematic reason why language should be so poor, but the fact remains that, for the discussion of contingencies of relationship, human language has yet evolved only a small vocabulary of words, which even the experts in human relations are unwilling to define in any critical manner. We refer here to such words as dependency, hostility, aggression, dominance, responsibility, spectatorship, prestige, respect, impertinence, rudeness, familiarity, intimacy, love, hate, and the like. For almost all of these words it is unclear whether they are descriptive of an individual, of the actions of an indi-

---

<sup>13</sup> For a discussion of many of these, see Bartlett (1987) and Bartlett (1992b). A very extensive bibliography of works relating to many varieties of reflexivity, comprising more than 1,200 citations, may be found in Suber (1987).

<sup>14</sup> Richard Routley took note of this fact from another perspective when he observed in connection with beliefs: "... since most intelligent animals have beliefs it is evident that no corresponding linguistic performance is required" (Routley & Routley, 1975, p. 230).

vidual or of a pattern of relationship.... We may suspect that these words still carry with them some of the ambiguity with which we endow them in our attempts to translate from analogic representations of contingencies of relationship into digital language. (Bateson & Jackson, 1964, p. 281)

Watzlawick, Beavin, and Jackson later went on to elaborate:

In human communication, objects—in the widest sense—can be referred to in two entirely different ways. They can either be represented by a likeness, such as a drawing, or they can be referred to by a name.... What then in analogic communication? The answer is relatively simple: it is virtually all non-verbal communication. This term, however, is deceptive, because it is often restricted to body movement only, to the behavior known as kinesics. We hold that the term must comprise posture, gesture, facial expression, voice inflection, the sequence, rhythm, and cadence of the words themselves, and any other nonverbal manifestation of which the organism is capable, as well as the communicational clues unfailingly present in any *context* in which an interaction takes place.... Indeed, wherever relationship is the central issue of communication, we find that digital [i.e., verbal] language is almost meaningless. (Watzlawick, Beavin, & Jackson, 1967, pp. 61-63).

The “analogic” or non-verbal varieties of reference tend to be overlooked in the restrictive focus of language analysts, where these varieties are often lumped together under the “ostensive reference” involved in the act of pointing, and yet they make up a major category of reference. They are legion in their varieties; another subset worth mentioning includes mental associations, in which memories, symbolizations, conceptual connections, etc., all play a role.

Still other forms of reference are to be found in the study of psychopathology, where we find, for example, delusions of reference (the belief that the actions of others have a special reference to oneself), delusions of influence (relating to social/contextual influences and individual psychological dispositions that predispose people to reason in certain ways so as to refer to evidence which they discount or reinterpret, and then favor certain alternative hypotheses), delusions of control (the person refers to someone else whom he

believes controlled his/her behavior), and delusions of grandeur, of persecution, and others. There are also delusions of misidentification, which include mistaken beliefs when referring to the identity of the person himself, other people, places, or objects. Such delusions are familiar to psychiatrists in the form of the Capgras syndrome (the belief that the patient's closest relatives have been replaced by imposters), intermetamorphosis (the belief that someone has changed physically and psychologically into another person), reverse intermetamorphosis (the belief that there has been a physical and psychological change of oneself into another person), and reduplicative paramnesia (the belief that there exist doubles of known people and places) (cf. Coltheart & Davies, 2000). Such psychopathological forms of reference have also been ignored by the mainstream philosophical study of reference.

My purpose in pointing to this great multiplicity of varieties of reference and self-reference is to underscore the fact that reference is ubiquitous in human experience, expression, and communication. The linguistic varieties of reference pale in number when considered in the context of these many shades and forms and pervasiveness of reference. Again, we reach the conclusion that there is no compelling evidence that we can obtain useful information about human thought and reality primarily, if at all, from an analysis of language in terms of which we describe thought and reality. We are led to the conclusions that the dominant, in-bred, constrained focus upon the analysis of language (a) rests on the baseless belief that a study of the structure of language will enable us to understand the nature and structures of thought and reality, and (b) is oblivious to the extensive variety of forms of reference that are not related to language. As a result, the view that a philosophical account of reference ultimately involves merely issues concerning language is a naive, excessively restrictive, and baseless belief which this book rejects. With that rejection we are forced to leave behind what mainstream philosophy of language has called 'theory of reference'.

### **3.4 From theory of reference to its metalogic**

We come to the main subject-matter of this book. It should be clear by now that this book, although it accidentally (i.e., contingently, non-necessarily) happens to find its expression in the English language and symbol set, will not be focused on symbols and their relations to other symbols and things other than symbols. Rather its threefold focus will be on the essential role of reference in many of our fundamental concepts and claims to knowledge, on ways in which that role is undermined, and on the self-undermining concepts, claims, and beliefs that we come to embrace as a result.

Thus far there have been two main alternatives to the study of reference: philosophy of language and psychology. The psychological approach, and here I include the phenomenological approach beginning with Brentano and culminating in Husserl, has sought to understand reference (a) in terms of intentionality—conceived as the relation that consciousness bears to objects of consciousness<sup>15</sup>—and (b) in terms of diverse psychological factors that are believed to play a role in establishing the reference relation, such as volition, agency, purposefulness, and the establishment of goals.

With the exception of the author's work, there has been no systematic, compelling attempt to go beyond these two alternative approaches so as to gain a level of comprehensive theoretical understanding of the most basic pre-conditions of reference, a level of understanding that is capable of self-application. Such a theoretical understanding establishes a framework of reference in which reflexive reference must be possible in order to account for that framework's own possibility. It is from the standpoint of this framework, to express this in very general and still undeveloped terms, that the meaning of the "metalogue of reference" will be developed.

When we attempt to reach such a comprehensive level of analysis that is applicable to the many distinguishable varieties of reference and self-reference, including such forms of reference as are needed to undertake this analysis, both the linguistic and the psychological approaches become excessively limiting, the first for the reasons given earlier, and the second because it carries with it heavy baggage consisting of philosophically and, in particular, epistemologically questionable ideas, beliefs, and propositions concerning human psychological functioning. This "baggage" routinely includes ideas, beliefs, and assertions that have to do, for example, with the commonly accepted distinctions between mind and body, between subjectivity and objectivity, and between consciousness and objects of consciousness, or that concern the presumption of the existence of a self and its allegedly active role in referring, the alleged "freedom" of such "acts of referring," etc. The "psychology of referential thinking" is laden with these extrania that distract and obstruct the investigator with entanglements in philosophically contentious issues which, from the highly general theoretical standpoint sought here, are themselves problematic and stand in need of the insights that a comprehensive theory of reference can provide.

To develop such a comprehensive theory of reference in the manner earlier chapters have affirmed—i.e., in a manner that meets the requirements of "rigorous science," of provability and incontestability—requires first that we

---

<sup>15</sup> Phenomenologist Marbach (1993, pp. 60, 175), for example, simply equates reference and intentionality.

spell out more specifically what this means, at the same time that we seek to divest ourselves of philosophical baggage that consists of baseless beliefs, and by so doing clear a way so we may proceed in a philosophically neutral, unbiased manner.

Einstein's definition of science will serve us as a starting point: "It would not be difficult to come to an agreement as to what we understand by science.... To put it boldly, it is the attempt at the posterior reconstruction of existence by the process of conceptualization" (Einstein, 1950, p. 24). Einstein likely had more implicitly in mind than he stated: To qualify as rigorously scientific means that the conceptual reconstruction to which Einstein refers is formulated so as apply *universally* to the phenomena that are its subject-matter; the underlying intent is to describe universal principles that govern or account for the subject-matter in view. And it must do this in a manner capable of *validation*, in physics, usually by empirical confirmation, or, as in mathematics, by abstract demonstration.

For a metalogical account of reference to qualify, from this point of view, as rigorously scientific, (1) it must provide a conceptual account of reference that cannot *not* be accepted without incoherence; (2) that account must be universally applicable to the many species of reference; and, because it comprises a non-empirical, purely theoretical undertaking, (3) such an account must be capable of theoretically based proof of its conclusions.



## 4

### **The Stepladder to Maximum Theoretical Generality**

Semantics, epistemology, ontology: any problem with roots in all three of these studies will be philosophically central and very difficult. Hence the tremendous interest in giving an account of reference. Reference has become the foundation of truth and thus crucial for semantics. Reference is our medium for speaking about the world, and therefore for formulating theories in which we express knowledge of the world. And since many ontological questions are settled (or at least debated) by appeal to the needs of semantics and epistemology, reference is a determinant of ontology as well. An account of reference will therefore have consequences for a broad variety of philosophical issues....

– Dale Gottlieb (1978, p. 79)

#### **4.1 Reference as the primary focus of this work**

**I**n the preceding chapter we saw how highly varied are forms of reference and self-reference. We have seen that one approach has been to conceptualize some of these varieties as relations between language or thought and the world; another has been to focus on the means by which objects are referred to. We have noted that reference, as we shall examine in more detail later, is often conceived as a function of relations between a speaker or thinker, a word or thought, an object, and a possible or real audience.<sup>16</sup> Some varieties of reference are verbal, some non-verbal, some attributed to psychological processes, and some are solely abstract, as in pure mathematics, bearing no explicit association with cognitive activity proprietary to the human species.

---

<sup>16</sup> For example, Bach (1987, p. 39) claims: “Following Strawson (1970/1950), I hold that reference is ultimately not a two-place relation between a word and an object but a four-place relation between a speaker, a word, an audience, and an object.”

To identify the most fundamental defining characteristics of reference and self-reference requires that we reach a level of universality and generality that can adequately encompass and throw light on their multiple varieties. Such a theoretically embracing objective should not be understood as an attempt to reduce the many varieties of reference to a single type, but rather to examine from a highly abstract level, which we will later characterize in terms of the concept of maximum theoretical generality, that which must necessarily be presupposed if these many varieties of reference are to be possible. It is on this purely abstract level that the focus of the metalogic of reference derives its meaning. Its purpose is to describe the invariant preconditions of all possible reference.

But why focus on *reference* (and, shortly, also on *self-reference*)? It has assuredly been one of the main defining objectives of philosophy to seek to reach a level of understanding that is so profoundly fundamental as to be universally applicable. To do this requires a highly refined degree of ultimate conceptual simplicity, of reduction to what is truly most basic, and in the process of realizing this aim, to attain a scope of breadth and comprehensiveness that is all-inclusive. These surely are difficult, elevated, ambitious, and some would say grandiose or simply impossible aims.

The growth of knowledge has already led to a degree of complexity that no one in a lifetime can take in and comprehend. At the same time, the more one delves deeper into a subject, the more one's knowledge and breadth of awareness become limited in scope. As the French phrase expresses this, the *idée directrice* of the metalogic of reference is to provide a conceptually powerful remedy for the resulting inherent unsatisfactoriness of overwhelming complexity and specialized narrowness. Its central objective is therefore the traditionally understood purpose of philosophy: to attain a unitary, unified grasp of reality and human experience on a level that is as genuinely fundamental, as conceptually basic, as it is possible for human beings to reach.

The study of formalized systems in mathematics has taught us that there is no such thing as a provably *most fundamental* set of axioms. The phrase 'most fundamental' and the ambiguous concept it expresses are of course a function of human purposes that define what the phrase and concept mean. We need to make those purposes explicit and as specific as possible, and in the resulting context develop a methodological approach that defines what we mean by a maximally fundamental understanding of that which underlies all human expression, communication, theorizing, belief, and knowledge.

We cannot make sense of our experience of physical, conceptual, artistic, or even psychopathological reality without recourse to some fundamental concept of reference. Whether that "sense" is reflective, communicable,

rational, aesthetic, or of another variety, reference in the broad meaning of that concept which we now have in view plays a conceptually and radically fundamental role. Reference is so basic to our conceptual vocabulary that its precise definition has been a philosophical challenge and, as one might expect, one that has not been solved in a single, definitive way.

Richard Robinson devoted much effort to understand the nature of definition itself; he observed in connection with the phenomenon of reference: "...the phrase 'is defined in terms of' can usually be profitably replaced by some such phrase as 'is referred to' or 'is located by reference to' (Robinson, 1962/1954, p. 180). The fact that underlies this observation, from the standpoint of the theory formulated in this book, is no accidental matter.

Reference, though it could conceivably be replaced by another equally fundamental concept that possesses the basic, comprehensive properties we seek, will serve in this work in a theoretically adequate capacity as a "most basic concept"—if for no other reason than that, in proposing any other concept that we might wish to substitute for it, we would be forced to make *reference* to such a concept. If one is compelled to make use of the concept of reference, and hence presuppose it, in order for it to be possible in principle to formulate an alternative most basic concept, then this "metallogical" fact confirms the logically primitive role of the concept of reference.

As we shall see, the concept of identification and of the identities of objects of reference will shortly also come to play a conceptually basic role so that the fundamentally abstract purpose in view can be realized. Because of the central role of identity and identification in the chapters that follow, I have for sometime thought of choosing the phrase '*metallogic of identification*' rather than 'metallogic of reference' to describe the nature of this study. My choice has been to continue to use the term 'reference', but in a metallogical sense, despite the word's different employment by linguistic theories of reference.

The reason for this choice is terminological appropriateness: The phrase 'metallogic of reference' has more intuitive direct and immediate applicability within a reflective framework that seeks to study phenomena, itself included, a framework which we are already accustomed to describe using the terms 'self-reference' or 'reflexivity'. Self-reference, or reflexivity, will be central to the development of the metallogic formulated here; the phrase 'metallogic of identification' fails to capture the close connection of this undertaking with what has become known as self-reference. With some ambivalence, then, 'metallogic of reference' is therefore the choice I've made in order to make clear that all varieties of reference, including self-reference, are taken into account, something which 'metallogic of identification' does not suggest.

With an interest in reaching an understanding of reference on a maximally abstract level, we stand in need of an explanation of the defining characteristics of this level of understanding. They are: universality, invariance, philosophical neutrality, the nature of the level of maximum theoretical generality itself, and transcendental analysis. We'll look at each of these in turn.

## 4.2 Universality

Universality is closely associated with degree of abstraction; generally, the higher the level of abstraction, the greater is the degree of universality that is reached. The adjective 'abstract' is a combined word derived from the Latin '*abs*' meaning "to drag away" and '*trahere*' meaning "to pull or draw." For a subject to be "abstract"—not merely in terms of its linguistic expression, but also in terms of its conceptual content—is for it to be disassociated, pulled away, from any individual, particularizing instance: In short, abstraction establishes a conceptual distance, a remoteness, from concrete, specific detail. This distancing of thought from the level of the concrete, from the level of individual things imbued with their richness of detail and specificity, is to translate thinking to a level which thereby has wider applicability because it is no longer tied to the level of the specific and concrete.

The noun 'abstract' means that a focus has been established upon the essential and more general characteristics of a group of specific things. The verb 'to abstract' then means that we have removed ourselves, in thought, from a consideration of specific things, and as a result have in view a broader, de-individualized view which it is appropriate to call more "universal." As the level of abstractness becomes greater, the scope of inclusiveness necessarily broadens so that what we have in view becomes less limited, less tied, less constrained by the specific instantiating, individualizing features of things on the concrete object level. Once we have "dragged away" and "pulled or drawn" from a collection of specific objects an abstracted set of properties, relations, defining qualities, etc., then that set is freed from particularity, and we reach a greatly widened scope of application and inclusiveness.

There are, to be sure, degrees of abstractness that can readily be distinguished: We have a familiar example in the traditional hierarchical International Code of Zoological Nomenclature, which, following the guidelines of the taxonomy of Linnaeus, classifies organisms, from the bottom up, by species, then genus, then family, order, class, and phylum in a nested sequence of progressively more inclusive categories which are, in a sense appropriate to taxonomy, increasingly more abstract and inclusive. (More recent competitive taxonomic systems such as cladistics and the PhyloCode, which use common

ancestry as a basis, do not exemplify a nested hierarchy paralleling orders of abstraction.) In descending levels of abstraction, as human beings we are members of the animal kingdom, phylum Chordata, subphylum Vertebrata, class Mammalia, order Primate, family Hominidae, genus Homo, species Homo sapiens, subspecies Homo sapiens sapiens. (As Isaac and Janet Asimov humorously commented, "...we named ourselves (doubly) sapiens, the most intelligent species. Nobody else calls us that" (Asimov, Janet & Isaac, 1987, p. 151).)

It is only a natural extension of this form of successively more abstract and inclusive thinking that we should conceive of a potentially top-most level in the hierarchy of theoretical abstraction. R. G. Collingwood called it 'the summit of the pyramid':

[W]hen the process of abstraction is pushed home to the limiting case and arrives at the summit of the pyramid, the thought which has effected this new abstraction might seem, therefore, to stand upon the threshold of a new science.... [It] stands in a situation not quite like the situations out of which ordinary sciences arise. The situation in which it stands is in certain important ways unprecedented and unique... (Collingwood, 1940, p. 13)

Collingwood had his doubts whether the "summit" of this pyramid of abstraction could result in scientific results, but he did not close the book on this possibility. Here, we intend to open that book's covers.

### 4.3 Invariance

In statistics, variance is a measure of difference with respect to what is most common to members of a group (the statistically inclined will recall it is the square of the standard deviation). We refer to the degree of variation in the heights and weights of individuals, for example. Speaking in more general terms, variation refers to the divergence in the characteristics of an object from what is typical or average in the group to which it belongs.

Invariance is a concept that functions in tandem with that of variance: It refers to what is constant and unchanging with respect to a group of objects. Invariance in physical theory and in mathematics refers to such constancy that is preserved despite physical or mathematical transformation operations.

When we ascend levels of abstraction, our interest is frequently in just such constancy that remains uninfluenced by individual variation. Often we

formulate the constancy that comes into view in terms of rules, principles, and laws that acquire theoretical and often practical value if they are in fact invariant over the particular types of objects studied. The formulation and verification of such general rules and principles define the core purpose of science and mathematics.

When we pair universality with invariance the focus of our attention resolves with increased clarity: Our attention is then drawn to highly abstract features of sets of things about which principles of regularity can be formulated. Should we be able to formulate such principles that are, as we shall see, in a very certain sense self-applicable, we will have ascended the rungs of the stepladder leading to what has been termed the ‘level of maximum theoretical generality’—that is, the “peak of the abstraction pyramid” hinted at by Collingwood.

#### 4.4 Philosophical neutrality

The concept of reference and its use in language are frequently qualified in a multitude of ways that call attention to psychological processes that we commonly believe are or must be involved when “we refer.” Reference to an object is customarily thought to be the causal result of an individual’s choice or intention; such choice is frequently believed to be “free,” and sometimes “involuntary.” In this way, the ordinary concept of reference is frequently encumbered—and by this I mean the basic concept of reference is laden with extraneous complexity—by beliefs concerning the “active role” of the individual person who “does” the referring. Here is some of the “baggage” I mentioned a few pages back, baggage that consists of questionable philosophical ideas, beliefs, and claims, often admitted and used without reflection, that carry with them a mass of issues that have weighed philosophers down for millennia in a quagmire of questions that have never yielded definitive, demonstrable answers.

The level of theoretical generality of the framework of the present study cannot afford to be encumbered in these ways, for only conceptually simple beginnings provide sufficient theoretical clarity and control over thinking and its expression. To rid ourselves of philosophical extrania is a preliminary methodological requirement that we can later decide with circumspection gradually to lift, once we have acquired a solid basis on which to stand. In what follows, the concept of reference will therefore be “philosophically neutralized” (some might say “philosophically emasculated”) in a way that leaves us with a theoretically purified concept of reference. There will be no talk or association of reference with “consciousness,” “subjectivity vs. objectivity,”

the “mind-body distinction,” “mind-dependence” or “mind-independence,” the presumption of a “self” that “does the referring,” or “the psychology of referential thinking.” Hence the concepts of intention, action, agency, volition, and their kin will be placed to one side, to be reconsidered later in the light of the metalogical principles of reference that will have been developed.

A theoretical framework committed to philosophical neutrality in the above sense adds a further condition of abstraction and broadens the universality that is the purpose of this general metatheory. It should be clear that ridding the study of reference of philosophical extrania—at least initially—is in keeping with the dissociation, the detachment, the “pulling away” from the level of the concrete, which, as we’ve seen, is ultimately what abstraction means. When Husserl set out to distance his phenomenology from “the naturalistic standpoint” and its “naive presumptions,” and formulated the idea of “bracketing,” he was motivated by much the same theoretical ideal, an ideal that is justified by the need to proceed rigorously, introducing a minimum of assumptions, and setting clear boundaries that define the scope of one’s theoretical interest.

There is a second, and more important, justification for the philosophical neutrality which the next chapters presuppose. That justification, in a few words, is this: We shall discover that most of the “philosophical baggage” that has occupied philosophers for so many centuries is thoroughly contaminated by a single widespread variety of provable referential error, whose elimination clarifies many of these problematic issues and shows them to involve forms of referential incoherence. When these issues are understood by means of concepts from which such incoherence has been eliminated, we find ourselves in a position from which we can develop a referentially coherent conceptual vocabulary that can serve as the rigorous, scientifically based approach urged in earlier chapters.

#### **4.5 The level of maximum theoretical generality**

And so we have before us an ideal goal: the establishment of a highly general, abstract framework, one that is philosophically neutral, maximally abstract, permitting what Einstein earlier considered a conceptual reconstruction of whatever subject-matter may be of interest, forming an adequate basis that makes it possible to describe governing or explanatory universal, invariant principles, and providing a way to accomplish this in a manner that is capable of a form of self-reflexive validation which cannot *not* be accepted without incoherence and which is universally applicable to the many varieties of reference. With this complex ideal goal before us, let us now begin to make steps

toward it.

In a very general sense, all philosophical positions are thought to bear the responsibility for accounting for themselves; if they are to do this, they need to be able to refer to themselves. As Collingwood noted: “Philosophy ... has this peculiarity that reflection upon it is part of itself.... [T]he theory of philosophy is itself a problem for philosophy; and not only a possible problem, but an inevitable one” (Collingwood, 1933, pp. 1-2). As Zaner explained:

Insofar as any philosophical position must be able to account for itself, ... it is itself constantly “at issue” in its own unfolding, and is by its very character compelled to take account of itself and its own possibility in its own terms.... This is not the case for other regions of human activity and thought: the theory of poetry is not itself poetry, but philosophy; that of science is not science, but philosophy; that of rhetoric is not a part of rhetoric, but philosophy. In a similar vein Husserl emphasized that if philosophy is to be at all rigorous, it must constantly and persistently make explicit and seek to justify *all presuppositions—the philosopher’s own most of all*. (Zaner, 1968, p. 75)

Gurwitsch developed this further, and implicitly connected philosophy’s responsibility to account for itself with self-reference:

Every philosophical system is subject to the obligation of accounting for its own possibility; it must at least be able to give such an account in its own terms. Less radically expressed, there must be no incompatibility between the doctrinal content of a philosophical theory, that which is maintained and asserted in it, on the one hand, and, on the other, the mere fact of the formulation of the theory in question. An incompatibility of such a kind would provide the basis for a decisive argument against the theory beset by that incompatibility. (Gurwitsch, 1966, p. 47)

Here we note parenthetically that detecting such internal incompatibilities in a philosophical position relies upon the possibility and justifiability of applying that position’s assertions to itself in a reflexive loop. The capacity of philosophical accounts to be applied reflexively to themselves has been widely employed by critics to bring out alleged or real internal self-referential



inconsistencies of this kind. The reflexive capacity of a theory formulated on the level of maximum theoretical generation will, as we shall see later, come to play an important role in this study.

Philosophers like Collingwood, Zaner, and Gurwitsch, articulate though they were, were not able to provide the clarity and explicitness that is possible from the more technical standpoint of formal logic. We owe to mathematical logician Frederic Brenton Fitch the first careful formulation of the concept, named by him, of maximum theoretical generality. Here is the way he expressed it:

Some theories are about theories. Others are not. Theories which do not include theories in their subject matter will be said to be of *ordinal level zero*. A theory which includes in its subject matter some theories of ordinal level zero, but none of higher ordinal level, will be said to be of *ordinal level one*. And so on. In general: A theory of ordinal level  $n + 1$  includes in its subject matter no theories of ordinal level greater than  $n$ , but it does include some of ordinal level  $n$ . Here  $n$  may be thought of as any finite or infinite ordinal number. Many theories proposed in the empirical sciences can be seen to be of some fairly low finite ordinal level. This is because empirical science is not generally concerned with framing theories about all theories.

A different situation prevails in philosophical research. Here extreme comprehensiveness is sought for. Theories are constructed which purport to deal with all entities whatsoever and which therefore have an unrestrictedly extensive subject matter. In dealing with all entities, such theories in particular deal with all theories, since theories are themselves entities of a special sort. In philosophy we thus encounter theories about the general nature of theories. If a theory has an ordinal level, its ordinal level must be greater than the ordinal levels of all theories occurring within its subject matter. Hence a theory about the general nature of theories can have no ordinal level, for its ordinal level would have to be greater than itself. Theories having no ordinal level will be said to be “vertical” or “non-ordinal” theories. Theories having ordinal levels will be said to be “horizontal” or “ordinal” theories.

If a theory is included in its own subject matter, we say that it is a self-referential theory. Since no ordinal level can

be assigned to a self-referential theory, every self-referential theory is vertical and non-ordinal. The converse, however, is not true, because a theory might contain vertical theories in its subject matter without containing itself in its subject matter. Such a theory would be vertical but not self-referential....

Any system of philosophy which takes a position on the nature of theories or propositions is itself a vertical self-referential theory. Particular views as to what constitutes a valid or acceptable theory are also themselves vertical self-referential theories. (Fitch, 1952 (a revised version of Fitch, 1946), pp. 217-219)

Fitch went on to say:

Although no *ordinal* level can be assigned to a theory which is about all theories, still we may speak of its “level” in some broader sense. A theory about all theories may be said to have attained the level of maximum theoretical generality. At such a level all other levels may be dealt with. There is no level which is higher in the sense that it can deal with theories not dealt with on the level of maximum theoretical generality. To deny that there is such a level is already to be proposing a theory about all theories and hence to be presenting a theory which is itself of the level of maximum theoretical generality. (Fitch, 1952, p. 223)

Fitch may have been the first to direct attention to and to name the level of maximum theoretical generality, but he was of course not the first to attempt to establish a philosophical framework on that level. Fitch mentions Whitehead's *Process and Reality* as an example of a theory about all theories, which includes itself in its own subject matter (Fitch, 1952, p. 218). Among others that come to mind, as Zaner noted, there is Husserl's theory of rigorous, transcendental phenomenology, which Husserl often called a ‘theory of theories’ or ‘science of science’. The metalogic of reference, as developed here, is another.

#### **4.6 Reflexivity on the level of maximum theoretical generality**

A self-referential, vertical theory capable of studying all theories establishes, then, a framework of reference on a level of maximum theoretical generality.

Such a theory, as Gurwitsch reminds us, is subject to a potentially lethal Achilles' heel in that the theory itself may not be consistent with claims it makes about theories in general. Such self-referential inconsistency on the part of a theory of maximum theoretical generality can of course be logically catastrophic.

Balancing this potential vulnerability, a self-referential, vertical theory can sometimes make possible reflexive proofs which demonstrate that certain claims cannot be denied without self-referential inconsistency. A special variety of such proofs will subsequently be called '*self-validating*'; they exhibit the compelling logical power that stems from universal claims formulated on the level of maximum theoretical generality, claims that cannot consistently be denied.

The metalogic of reference formulated in this book is a vertical, non-ordinal, reflexive theory developed on the level of maximum theoretical generality. But unlike traditional vertical, non-ordinal philosophical theories, the metalogic of reference explicitly studies the concept of reference as its core subject-matter, which, as remarked earlier, possesses the conceptually fundamental status of being logically primitive, that is, reference must be presupposed by any theory. Recognizing that reference in this sense is a "most basic concept," without which we cannot think, speak, or formulate *any* theory (even a theory that might be an alternative to the present one), a theory of reference developed on the level of maximum theoretical generality will itself inescapably necessitate a structure capable of reflexivity, of self-reference.

Fitch recognized the intimate connection between a theory articulated on the level of maximum theoretical generality and the ability of that theory to allow for self-reference. The system of logic that he developed (Fitch, 1952) was formulated precisely in order to "find a logic which eliminates the 'vicious' sorts of self-reference that lead to the mathematical and semantical paradoxes but not those sorts of self-reference that seem to be such an important part of philosophic logic" (Fitch, 1952, p. 225). By "philosophic logic" Fitch meant the logic, in his view, that is required by any philosophical theory that is able reflexively to account for itself.

#### **4.7 Expanding the scope of a maximally general theory of theories in order to study preconditions of possibility**

It is one thing for a philosophical position to bear the responsibility to account for itself, and another for it to respond to the need to account for its very possibility. At times these two tasks are lumped together, but they are distinct. In the first case, such self-accounting requires a reflexive analysis and expression of the philosopher's most basic assumptions, definitions of fundamental

concepts, clarification of the range of subject-matter to which his or her propositions apply, etc. In the second case, such self-accounting requires a more generalized analysis of the preconditions of any inquiry that is of the sort exemplified by the particular philosophical position in question. Preconditions of this kind are more abstract and logically fundamental than is the first form of philosophical self-accounting. They often relate to presuppositions that must be made and presupposed concepts that must be available in order for any such a framework to be possible in principle. Philosophical accounts that employ transcendental arguments can be of this kind, of which Kant's account is a classical example.

#### 4.8 The concept of metalogic

Two approaches to the study of maximum theoretical generality need at this point to be distinguished: Fitch identified the first, that of a general theory of theories; the second relates to the method of analysis of maximum theoretical generality proposed by the metalogic developed here.

The term 'metalogic' has traditionally been used, for example by Carnap (1959/1932), to refer widely both to the theory of expressions of a language, usually a formalized language, and to the logical relations existing among those expressions. Once the consistency and completeness of systems of propositional calculus began to be investigated as objects of study, as, for instance, by Łukasiewicz, or by Tarski in his formalized metalogic in which the objects studied are entire deductive systems, the study of metalogic developed a more specialized meaning, which is properly a part of metamathematics. Metalogic in the sense of Łukasiewicz or Tarski seeks to formulate general principles that govern the systems of logic that are in view, and hence are subordinate to the meta-level upon which analysis occurs.

When we establish a superordinate frame of reference that permits reference to subordinate specific systems or theories, or to systems or theories in general, the superordinate reference frame is often described as a metalanguage, and the formal study of the subordinate systems is called 'metamathematics' or 'metalogic'.<sup>17</sup>

The metalogic of reference developed here is formulated from the standpoint of a "transcendental" approach to maximum theoretical generality, and comprises a "metalogic" in the sense that it formulates general principles which govern possible reference in any subordinate theory, general principles which also govern possible reference from its own standpoint. The metalogic

---

<sup>17</sup> See, for example, Feys and Fitch (1969, ¶ 90.1).

of reference formulated in this book could therefore be called a ‘transcendental metalogic of reference’, but for economy the term ‘transcendental’ will be dropped. A full account of the concepts of “transcendental” and “metalogic” will be developed as we proceed.

#### 4.9 Meaning

A plethora of theories of meaning is to be found in the history of philosophy, in the history of linguistics and semantical theory, and of course in the more humanistic, less precisely formulated accounts of meaning in such fields as literary criticism, aesthetics, and psychology. It would take this book far afield to discuss and critique even the main philosophical theories of meaning. The reader will by now have realized that the theoretical framework that will be in view in this study is highly abstract and intentionally dissociated from particularizing, “material,” content. In order to achieve the level of inclusiveness and generality sought, the meaning of ‘meaning’ from this highly theoretical standpoint will have an equally abstract, broad scope.

Philosophers commonly distinguish two main varieties of meaning, often called ‘cognitive’ as opposed to ‘expressive’ or ‘affective meaning’. Cognitive meaning is frequently understood as the kind of meaning that referring statements, propositions, and concepts are thought to have, as distinguished from the psychological, emotionally based meaning which works of art, literature, or a person one loves may have.

Cognitive meaning has traditionally defined the boundaries of applicability of linguistic theory of reference, which has most often placed the affective variety of meaning to one side to be studied in the looser and often vaguer styles characteristic of less formal and less technical humanistic approaches. However, to do this loses sight of the fact that reference plays an equally important role in connection with both forms of meaning. As a result, the cognitive-affective distinction will not be considered useful here. Instead, meaning itself—if a unitary concept descriptive of all forms of meaning could be articulated—is left open and undefined, although a necessary precondition of meaningfulness will be defined later on. That precondition is referential consistency, which ranks as a most fundamental criterion of meaning. In applying that criterion, any putative cognitively meaningful statement, proposition, or concept, or any apparent reference to emotionally based meaning, which is such that its referential structure is inconsistent with itself—causing it, so to speak, to undermine itself or implode—will, as we shall see, be considered, from the reflective standpoint of metalogical analysis, to be devoid of meaning.

The purpose of this study, then, is the development, self-validation, and application of this generalized criterion of meaning, one which will enable us to identify, and then to eliminate, certain well-defined forms of self-referential inconsistency, and by doing this, will lead to a recognition, one that cannot *not* be accepted without referential incoherence, of the most general, invariant principles of reference that govern what can—in principle—be meaningful.

#### 4.10 The self-enclosure of reflexive, maximally general theory

Among the fascinating objects that can be studied in topology, we encounter some that form two-dimensional surfaces or three-dimensional spaces that curve back on themselves. The Möbius strip is a two-dimensional surface, a model of which is easily made by giving a paper strip a half-twist, then connecting the ends to form a loop. A caterpillar crawling along the length of the strip will be able to move along the entire surface of the strip without leaving the surface and without crawling over an edge, proof that the strip in fact has only a single continuous side. As long as the caterpillar remains on the surface, it remains confined to a two-dimensional universe that has the property that I shall call '*self-enclosure*'.

Similarly, a Klein surface, which has become known as a Klein bottle, recurves on itself. Although three-dimensional suggestive models of a Klein bottle have been made, to construct a genuine Klein bottle requires a fourth spatial dimension, whose physical reality topologists have yet to bring about, though many would surely like to! Where our caterpillar can encounter an edge as it crawls along a Möbius strip, a Klein surface has no edge; it forms a continuous self-enclosing surface. Both the Möbius strip and the Klein bottle exemplify the property of self-enclosure.

In much the same way, self-enclosed is the topology of the space-time manifold of a relativistically recurved physical universe: In such a universe, a space traveler will never encounter a physical boundary beyond which lies a spatial "outside" or a temporal "before" or "after." Instead, if the universe's topology has, for example, a spherical metric, the space traveler, like a caterpillar crawling on the surface of a Möbius strip or Klein bottle, will be able to return to his point of departure if he travels far enough. The essence of self-enclosure is the absence of boundaries in a manner which, despite their absence, results in a closed system.

Philosophical systems have the property of being self-enclosed in a similar and non-metaphorical sense. As Gryn timer (1961) remarked, they are developed specifically so as to apply to the object of their study, and as a result form "*ensembles fermés sur eux-mêmes*," which is to say, "sets that are closed

upon themselves,” or, as I express this in what follows, are “self-enclosed.” We have already encountered the philosophical variety of self-enclosure in connection with the closed nature of some systems of belief: They confine the believer to a point of view that he cannot/will not go beyond. No matter what challenging information may be presented to him, he will remain “recalcitrant,” re-interpret all so as to confirm his beliefs, and stubbornly remain unconvinced in the face of all evidence to the contrary. I have called this form of belief recalcitrance a kind of intellectual narcissism, which succeeds in keeping its possessor in a hermetic, self-enclosed system of belief.<sup>18</sup>

Self-enclosure, then, may be topological, or systemic as in the case of a set of beliefs. Self-enclosure also characterizes a reflexive, vertical, non-ordinal frame of reference established on the level of transcendental, maximum theoretical generality. From such a standpoint, a “universe” comprehending all possible reference is defined; reference “outside” such a framework of reference is impossible since any reference will, by definition, form part of that universe.

There are, in addition, expressions of self-enclosure that are especially important to philosophers and psychologists: For phenomenologists, experience comprises a self-enclosed “field”: No matter how experience is extended, it is never possible to “go beyond” its boundaries, for though bounded, it has no “boundaries” as these are conventionally understood; whatever is experienced becomes part of the field of experience. For phenomenologists, this fact is a non-trivial tautology: No matter the experience, it remains part of the self-enclosed field of experience. (For phenomenologists, the previous two sentences can be re-stated by replacing ‘experience’ by ‘consciousness’ or ‘awareness’.) Elsewhere, I have studied what I termed the ‘logic of structure’ of this phenomenologically recurved space-time continuum that we call ‘experience’.<sup>19</sup>

In much the same way, the systems of belief of delusional psychiatric patients can be no less refractory in the face of opposing evidence, and can possess a recalcitrant hardihood sufficient to challenge any clinician. Some of the varieties of delusional reference were mentioned in {3.3}; all exhibit self-enclosure.

The property of self-enclosure will later prove to be fundamental to the capacity of the metalogic of reference, as it is developed here, to establish

---

<sup>18</sup> Bartlett (1986a).

<sup>19</sup> Bartlett (1970).

universal, invariant principles of reference that cannot *not* be accepted. We shall find that self-enclosure also determines many of the limitations of reference that govern what we can claim meaningfully to know.



*A horizon defines—from your present standpoint—not only how far in fact you see, but how far you can possibly see.*



# PART II

## THE METALOGIC OF REFERENCE

### A New Approach to Deductive, Transcendental Philosophy

This part of the book describes theoretically fundamental principles without which referential consistency, meaning, and hence coherent, intelligible experience become *impossible*. The proof that a given principle is a “precondition of possible reference, meaning, and intelligible experience” might be thought of as a contemporary form of “transcendental deduction,” the logical structure of which will be made clear in what follows. Roughly speaking, each candidate for the role of a precondition of possible reference is tested by attempted denial. If such a denial results in metalogical self-referential inconsistency, the principle’s role as precondition is confirmed, otherwise it is rejected.



# 5

## Reference, Identification, and Identity

### 5.1 What is reference?

The phrase ‘*to refer to*’ functions in ordinary use as a means to direct attention to something—in other words, *to point*. When used in this attention-directing way, we call this ‘ostensive reference’. Pure ostensive reference happens just by using a finger to point; using language ostensively is a small but important further step. Using language to point is probably one of the most ancient uses of language, developed to enable one person to communicate a wish for another person to become aware of an object to which the first person desires to call shared attention.

As a result of this basic purpose for which language is often employed, it has felt intuitively obvious to the majority of philosophers, as it does to non-philosophers, that referring is an essentially relational concept. In this customary and conventional view, and with language use explicitly in view, referring has been conceptualized as linking a symbol, word, or phrase with an object, thereby establishing a relation between the language that is employed by the speaker who refers, and, as a result, establishing for a hearer relationships among the language the speaker uses, the object that is pointed at, and the speaker who is doing the referring, and by this process permitting an exchange of information between speaker and hearer that we call ‘communication’.

In common use, the term ‘relation’ is used to claim that there is a connection or a contrast between things that differ from one another. As John Locke (1690, p. 151) expressed this, “[t]he nature ... of Relation consists in the referring, or comparing two things, one to another.” When the claim is made that there is a relation between two things, we call these ‘relata’, and as Stebbing (1933/1930, p. 111) put this, “[t]he term *from* which the relation proceeds is called the referent; the term *to* which it proceeds is called the relatum.” When in the present study, I use the term ‘non-relational’ it is to de-emphasize, or to avoid entirely, the notion that distinguishable relata are involved.

The basic relational conception of reference is, surprisingly, seldom taken

apart and examined in detail, even by philosophers who study theory of reference. Playing important roles in the “active referential process” that is believed to be involved in ostensive reference are implicit or implied claims about what goes on when there is reference to an object. Usually a number of such claims play a part, claims which on the surface look innocent enough, but which have been responsible for unnecessary confusion and, as we shall see, lead to conceptual mistakes that have had widespread consequences in the thought of philosophers and even in the everyday thinking of non-philosophers.

In this work I distinguish two fundamentally different concepts of reference. The first, the conventional and now well-established, is the *pragmatical concept of reference*, which has come to dominate language-centered studies in the theory of reference. In labeling this view “pragmatical” I follow the threefold division of semiotics by Charles Morris,<sup>20</sup> according to the ways in which symbols can enter into different kinds of relations: *syntax*<sup>21</sup> pertaining to relationships among signs, symbols, and sentences formed of these; *semantics* dealing with relationships between such signs, symbols, and sentences and the objects they signify; and *pragmatics* extending the foregoing relationships to take into account language use by persons.

In the following section, I discuss the pragmatical concept of reference, and later turn to the second and conceptually more fundamental *metalogical* concept of reference.

## 5.2 Pragmatical reference

As we have seen, since the middle of the twentieth century, reference, as a subject-matter studied in itself, has become the nearly exclusive territorial province of philosophy of language. And as has previously been touched upon, philosophers of language have understood reference in terms of various kinds of relations that may occur between linguistic symbols and other symbols, or between symbols and things other than symbols. These language-based relationships they have discussed and analyzed under such familiar headings as denotation and meaning, and the extended family of their cousins that have been distinguished in the history of semantics.

Although accidentally (i.e., contingently, non-necessarily) expressed in the English language and its alphanumeric symbol set, this book, in contrast with philosophy of language, studies not symbols and their relations to other

---

<sup>20</sup> Morris (1946).

<sup>21</sup> His term was ‘syntactics’; ‘syntax’ is now commonly used instead.

symbols and to things other than symbols, but rather the focus is a set of concepts; many of these are concepts that we routinely use and misuse, whether or not we are philosophers; and related to this conceptual focus are the beliefs that people come to embrace as a result.

In this context, what I shall mean by ‘reference’ is an abstract concept that may best be understood by contrasting it to what I will call ‘the naive view of reference’, which has become well-established, habitual, and virtually standardized, comprising the dominant dogma in the language-based literature about reference. The following are defining claims that make up the basic ingredients of the naive conception of reference:

- (1) reference is an *activity*, expressed by people through their “*performative linguistic acts*”;
- (2) it is, moreover, a *volitional* activity, one that expresses the purposive *intentions* of human *agents*;
- (3) this performative activity of referring takes place in *time*;
- (4) it is usually, but not necessarily, presumed to be an activity that takes place “*freely*,” in the sense that it is believed to be free of prior causal determinants;
- (5) the human agent’s referring intentions may be “*successful*” or not, depending on whether the agent’s linguistic referring activity targets his *intended object(s) of reference*;
- (6) referring activity takes place through the use of various “*tools of reference*,” that is to say, linguistic means or devices, consisting of words, phrases, or other symbols or signs, which may be descriptive, or proper names, or other expressive locutions;
- (7) the agent usually, but not always, intends to refer so as to *communicate* to others what he intends, and hence there is often an *implied audience* that plays a role in his referring activity; and finally,
- (8) all of the above most often take place within a concrete “*context of reference*” that situates the agent, provides the backdrop for his referring intentions, the temporal occasion for the linguistic means that he uses to express these intentions, the range of possible objects to which he may intend reference, other parties to whom his referring intentions are to be communicated, and, as a function of all of these, this process takes place in a manner that permits some determination to be made of the “*success*” or “*failure*” of the agent’s referring intentions.

These claims that make up the naive view of reference are pervasive in the language-focused literature devoted to theory of reference, and have become dominant from the time of the early publications by Strawson to the present.<sup>22</sup>

Now and again, some philosophers of language have made it clear that they have in view what ultimately are psychologically based human cognitive *abilities*, as when Stephen Schiffer equates theory of reference with a theory of thought,<sup>23</sup> or when Timothy Williamson automatically associates reference with a subject's "mental states" and "acts."<sup>24</sup> For such philosophers, referring is an ability that people have, an ability that these authors choose to study primarily in connection with the human ability to use language.

As we can see in (1)–(8) above, an assortment of very diverse ideas has been compounded in the now conventional account of linguistic reference. Before we proceed, it is important to re-emphasize this fact and attempt to gain a measure of dispassionate distance from it: The philosophical understanding of reference that has resulted has come, with comparatively little critical self-reflection, to embody an account of individual human functioning that is conceptualized in terms of the referrer's mental states, volitional acts, purposive activity, and his or her use of referring linguistic means, with all of the preceding believed to express mental abilities whose exercise, in relation to a presumed or intended context, may be publicly judged successful or not in an interpersonal context of communication. The naive view of reference is therefore an amalgamation of a wide variety of explicit, implicit, or implied claims which seldom are specifically identified, brought to light, and examined critically by philosophers of language. As noted earlier, these claims very evidently bring along with them the ponderous weight of philosophically mixed and unexamined baggage.

The baggage accompanying the concept of reference has unavoidably made a clear and effective theoretical understanding of reference both difficult and problematic. For each of the above eight claims comprises, in and of itself, a traditionally debatable topic of philosophical analysis, ranging from the application of some presumed conception of the self or agent, its volitional

---

<sup>22</sup> For example, Strawson's often-cited paper, "On Referring" (Strawson, 1970/1950), explicitly asserts (1), (3), (4), (6), and (8), and implicitly appears to endorse the other three. His later, also often-cited paper, "Identifying Reference and Truth-values" (Strawson, 1970b/1964), reaffirms these commitments and makes evident that he has in view a human referring *ability* (the ability "to pick a thing out," p. 98) that is basic to communication.

<sup>23</sup> "The basis of a theory of reference must ... be a theory of the thought in the mind of a person using a singular term; typically, this thought is a thought about the object referred to by the singular term on the particular occasion of use. So the basis of reference is a theory of our thoughts about things..." (Schiffer, 1978, p. 171).

<sup>24</sup> "Reference concerns what mental states and acts are about" (Williamson, 2007, p. 269).



capacity, perhaps its freedom of will, its linguistic action directed toward intended ends, embedded in an assumed temporal context in which there is an implicit separation, on the one hand, between the referring linguistic means employed by the referring agent, as distinguished, on the other hand, from the object of reference intended, and situated in a presumed intersubjective world in which the referring mind is related to other minds.

To develop a philosophically neutral, maximally general study of reference, it should be evident that we need to resist casually importing these many unanalyzed preconceptions, however intuitive and habitual their use has become. Resisting this temptation will in the end turn out to be a wise decision as it becomes clear that many of the individual claims that are ingredient in the naive view must compel the rational person to reject them on grounds of incoherence, and to replace those that can be salvaged with a conceptually sound metalogical understanding of reference.

Before we leave this preliminary summary of the naive view of reference, two of its conventionally propounded claims stand in need of more detailed attention.

### **5.3 The separation of reference from the object of reference**

At the basis of the naive view of reference is the traditional separation of referring term and object referred to. This separation has likely occurred as a consequence of two general ways in which people communicate about objects: To direct attention to a certain object, the object can be depicted, for example, by means of some kind of visual representation, a drawing or photograph, or the object can be described or named by linguistic means. The picture is not the object, nor is the linguistic expression the object. It is natural to recognize the difference.

This separation of the object from the means of directing attention to it has been comfortably and usually unquestioningly accommodated in the philosophically unreflective everyday view of the world in which it feels natural and intuitive to separate minds from things, acts of pointing from things pointed at, etc. Although these separations have become habitual, habit does not imply that they are necessary or that they comprise the only possible or most effective way of conceptualizing what most fundamentally is involved. Especially is this true should the naive view of reference suffer from forms of conceptual incoherence.

For the present, it will be enough if we only place a question mark after the conventional separation of reference and referent, and leave the matter

unresolved until a later point in this study when it can be considered using the technical means that will then be at our disposal.

#### 5.4 Removing agency, intention, action, and volition—at least provisionally—from theory of reference

As we shall see, it unnecessarily complicates and confuses the referential rules that will be our object of analysis were we to feel a need always to write or otherwise express ‘ $\text{↗} p$ ’ instead of, simply, ‘ $p$ ’. The pointing finger symbol is intended to embody or connote what is commonly believed to be the ostensive reference, the “targeting of attention” which the naive view of reference claims is involved.

In order to avoid introducing concepts that we shall later find bring their own special problems with them, it can be a challenge and certainly contrary to habitual use to develop ways in which we can employ ordinary language, but in a manner so as to avoid customary modes of expression that carry with them those very problematic issues. Like many languages, the English language, as it has so far evolved, is inclined, as we’ve noticed, by reason of its users’ habits and by its grammar, to associate “acts” and “agency” with referring statements. Because its purpose is to bring the user of language into the explicit scope of analytical study, the pragmatist view of referring is permeated with terms that implicitly or explicitly characterize the role of the user of referring expressions. As a result of this interest, John Austin,<sup>25</sup> for instance, distinguished “locutionary acts,” “perlocutionary acts,” and “illocutionary acts.”—For example: I exclaim, “It’s hot in here” (a locutionary act), which expresses my sensitivity to heat (illocutionary—what has been done in the very saying of what I’ve expressed), and as a result my wife turns on the air-conditioning (perlocutionary—what is accomplished by what I said). Since Austin’s time, finer and more specific distinctions among linguistic acts have multiplied many times over under the rubric of speech act theory.

Proceeding in this way, the study of reference has become very strongly activity-centered. Users of a language are described as variously engaged in *actively referring*, suggesting that reference involves multiple psychological processes and their effects. Even the language chosen to talk about reference in a de-personalized way can inherit this idiomatic manner of speaking: For example, “an expression *makes* reference to a certain object of reference,” although expressions of course do not actively engage in or manufacture anything.

---

<sup>25</sup> Austin (1962/1955).

So ubiquitous and second-nature has this manner become of thinking and of expressing statements about reference that it may be hard to imagine how else reference could or ought to be understood. And yet if we wish, in an effort to reach a maximally general theory of reference, to remove the swarm of potentially extraneous and unexamined elements that have come to surround the use of the verb ‘to refer to’, we shall need to find a more neutral way to conceptualize reference.

### 5.5 Levels of reference and iterative reference

In connection with some object,  $a$ , we can distinguish  $a$  from the means,  $\text{☞}$ , used to point to  $a$ . We could further make  $\text{☞}$  the object of attention by expressing  $\text{☞☞}$ . This manner of expression is broader than the autonomous use convention of placing inverted commas around a symbol, name, or expression in order to refer to it, as in “The word ‘four’ consists of four letters, whereas ‘4’ consists of only one digit.” We could accomplish this, too, by using ‘☞’ (enclosed in inverted commas) to make the pointing finger symbol itself the object of attention, that is,  $\text{☞ ‘☞’}$ . The autonomous use convention was designed to allow unambiguous reference to symbols, names, or expressions, whereas the pointing finger symbol can be used in a less restricted way to point to an object,  $\text{☞}a$ , or to point to the pointing of that object,  $\text{☞☞}a$ . Symbolized more conveniently and less pictorially, we can distinguish  $a$  from  $Ra$  and from  $RRa$ , a manner of symbolic representation that I shall use in what follows.

Clearly, references can continue to be iterated. With each iteration, a different object of reference is specified and with it a different level of reference is defined. The properties of object  $a$  may be characterized in various ways; a description of  $Ra$  (requiring  $RRa$ ) no longer has  $a$  as its object, but rather  $Ra$ ; etc.

Reference can in principle be iterated indefinitely, though soon this may strain our cognitive abilities to keep clearly in mind what we are to take to be the object that is at issue. This is additive iteration. But an iterated reference can also be reduced, as when a transition is made, for example, from (1)  $RRa$  to  $Ra$ , or from (2)  $Ra$  to  $a$ . The naive way of expressing this fact is to say that in (1), we “intend” to talk about  $Ra$  and not about  $RRa$ , whereas in (2), we “intend” to talk about the object  $a$  and not about  $Ra$ .

Consider a specific context of reference—say, when working with a particular coordinate point  $(1,0,+\sqrt{2})$ . A frame of reference is presupposed, conventionally a Cartesian coordinate system with three axes, which makes it possible to locate the specified coordinate point. The foregoing sentence that I

have just written involves a “stepping-back” from the framework that enables us to make explicit how the identity of the particular coordinate point is to be understood, that is, in terms of a Cartesian coordinate system; this “stepping back” is an example of  $Ra$ , where  $a$  is that coordinate point. (And this last observation is an example of  $RRa$ , a further “framework step-back.”)

Often, a framework step-back makes explicit what was taken for granted implicitly one level of reference lower.

An ability to keep such levels of reference clearly distinguished will prove to be fundamental to the metalogic of reference that will concern us.

### 5.6 A non-linguistic, non-relational conception of reference

To reach a level of maximum theoretical generality concerning reference, there is, to express this concisely, at least a provisional need to strip away from the standard concept of reference the group of eight claims identified a few pages back. The traditional use of the term ‘reference’ and its permitted range of application have lumped these claims together, leaving them fundamentally unanalyzed from the metalogical standpoint we shall later explore, a standpoint which will make it possible to examine the referential preconditions that must be granted in order for such claims themselves to be possible. This highly general and abstract form of analysis will occupy us in later chapters; but at this point, the naive view of reference already can be depicted in the following more explicit fashion:

Beginning with some specific object, a proponent of the naive view engages in a framework step-back in directing attention to the object’s relation to a linguistic sign used to refer to it; a further step-back often then points to the speaker’s cognitive intention or to his or her perceptual state or to other properties of referring agents; a further step-back broadens the range of reference to include perhaps other people who comprise the speaker’s audience; etc.

Understood in this manner, the “stripping away” of the extraneous claims of the naive view amounts, then, to a decision to reduce that view’s built-in and implied iterated references so as to point simply to objects of reference, which then will comprise the initial scope of reference that we wish to study in some detail.

Stated somewhat differently: The traditional conception of reference and of referring makes a separation between object and sign or symbol used to refer to that object, but here, from a metalogical standpoint, reference is fundamentally identify-specification or identity-recognition, so that the object and its identity are not disjunct, but essentially fused: there *is* no object with-

out its specifiable identity. This is what is meant by a “non-relational” conception of reference.

Immediately, the question should suggest itself to the reader: Can reference be understood in such a purely non-relational sense? what would the result be?

Expressed intuitively and impressionistically, the result *collapses reference upon itself*, divesting reference from its multitude of automatic associations, from habitual uses deeply ingrained by the styles of expression that present-day English and similarly structured languages encourage, and from the equally deeply entrenched ways in which we have become accustomed to thinking about reference. We have become completely inured—habituated to the point of obliviousness—to what we take to be self-evident agency-based claims such as “there can be no reference without a referring agent,” “no thought without a thinker,” “no perception without a perceiver,” “no consciousness without a self,” etc. To go beyond these, even in a provisional way as is suggested here, will be felt by many as a strange or uncomfortable challenge, certainly one that is counterintuitive. These are natural reactions when habitual linguistic presumptions are called into question.

In addition to linguistic reference, non-linguistic reference is commonplace in our lives: from pointing fingers, to gestures, to situationally determined reactions, to artistic representations and musical expression. There is no end to the great variety of non-linguistic examples of reference. In their diverse, often very distinct, ways of specifying, recognizing, expressing, or communicating, they share in common the capacity to express, depict, or represent, sometimes ambiguously, sometimes precisely, some object or group of objects—and here the word ‘object’ is deliberately employed very loosely, to apply to what may be nebulous, fuzzy, indeterminate; qualities or feelings; specific, concrete physical states, things, or measurements; or purely abstract things—numbers, shapes, concepts, properties of formal systems; etc. Sometimes such objects are of course identified as referents of natural or artificial language expressions, but frequently such symbol-based language plays no part.

The conceptual core of all of these varieties of reference is *identification*, identification that is taken in a broad and inclusive sense. Whenever we can, to some significant degree, “understand what we’re talking about, what we’re expressing, referring to, depicting, or studying,” some degree of identification is necessarily involved. Stripped of its pragmatical associations, its habitual linkage with linguistic acts and agents, the conception of reference that we are led to is embedded in identification that is not explicitly itself relational in nature. A non-relational conception of reference, as will be developed, is a

“reduced” or “minimalist” conception of reference, in which reference is intimately tied to whatever framework of reference is necessarily presupposed in order for such reference to be possible, but such a conception of reference refrains from endorsing the eight claims of the naive account of reference listed earlier in this chapter.

### 5.7 Identification, reference, and coordinate systems

Much like the intransitive verb ‘to refer to’ and its related noun ‘reference’, the verb ‘to identify’ and its related noun ‘identification’ suffer from the same eight pragmatical presumptions listed toward the beginning of this chapter. It should not be necessary to retrace the foregoing discussion as it applies to the latter terms; instead, we begin by considering a stripped-away conception of identification, one not linked to linguistic agents and their acts.

As I elaborate in more detail in the next section, reference is essentially tied to identification: any instance of reference is at once an instance of identification, whether such identification is vague or precise. Any object to which reference is possible *is* an object that possesses an identity. To be a possible object of reference is, in other words, to be an entity with some degree of identifiable identity. The point  $(1,0,+\sqrt{2})$  can be referred to precisely because it is identifiable. Its identifiability and its identity as a coordinate point are a function of the rectangular Cartesian coordinate system that is the presupposed basis for its identity and identifiability.

Coordinate systems are normally associated with mathematics and with their application in science. This association is unnecessarily restrictive: Frames of reference make it possible to refer to diverse ranges of objects, and thereby to recognize the identities of those objects (to know what is specified, described, adumbrated, etc., as distinct from all that is not so specified). Frames of reference function as systems of coordination in real and non-metaphorical ways, though not all frames of reference have well-defined axes permitting numerical coordinates. Frames of reference function as systems that permit the coordination, for example, of locations with objects; of persons in relation to things; of things and certain times; of persons, things, and times, etc. Such “ coordinations ” are usually presumed implicitly, as part of the contextual background, but they are, as we shall have occasion to develop in some detail later on, the very basis for our ability to know and to communicate knowledge of any class of objects of reference.

Few philosophers who have occupied themselves with the theory of reference have recognized both the fundamental role of the coordinative function of frames of reference, and the explanatory value of coordinate systems in the

theory of knowledge and in the theory of reference. Occasionally a nod of recognition is given, but it then tends to be passed over quickly as though the coordinative function of reference and its role in knowing are obvious and stand in need of no explicit study.

One of the earliest and most important serious attempts to apply the mathematical concept of coordinate system within a philosophical study was a book published by Hans Reichenbach a century ago,<sup>26</sup> four years after Einstein published his general theory of relativity (a theory in which attention to the role of coordinate systems is central, as we shall examine later in {26}). Reichenbach's purpose in his book was to develop, in a Kantian framework, an epistemological understanding of the concept of physical object as it is found in Einstein's special and general theories of relativity. He considered Kant's transcendental theory of knowledge to provide the most effective epistemological frame of reference in terms of which to understand relativity physics: "I believe that this theory [Kant's theory of knowledge] stands unexcelled by any other philosophy and that only it, in its precisely constructed system, is *equivalent* to Einstein's theory [of relativity] in the sense that a fruitful discussion can ensue" (Reichenbach 1965/1920, p. 112, n. 17, italics added).

Although Reichenbach's slim book, really a monograph of just over a hundred pages, does not provide justification for his central claim above, the book is useful as background for our later discussion and will be reviewed briefly here.

Given his specific interest in theoretical physics, Reichenbach's application of the concept of coordinate system was intended by him, in an essential theoretically basic way, to relate physical observations to the reference frames that make them possible. "The physical relation [to reality] can be conceived as a coordination: physical things are coordinated to equations" (pp. 36-37). "[T]he 'real' is defined by coordinations to the equations" (p. 38). In this relativistic framework, Reichenbach's concept of *coordination* functions as "that most general principle ... [that is] presupposed by all knowledge... [C]oordination seems to us to be the most general concept that describes the relation between concepts and reality" (p. 86).

Reichenbach shared Kant's transcendental interest in identifying the preconditions necessary for the possibility of objective knowledge; in Reichen-

---

<sup>26</sup> Reichenbach (1965/1920). This work, originally published in German, was not translated into English and published until 45 years later, in 1965. (N. B.: A central chapter in the book, Chapter IV, is listed in the book's table of contents as "Knowledge as Coordination" (pp. 34-47), but the chapter itself and all headers of pages in the chapter bear the title "Cognition as Coordination.")

bach's case, this became an interest in identifying what he called 'principles of coordination': "principles of coordination ... define the individual elements of reality and in this *constitute* the real object" (p. 53).<sup>27</sup> Principles of coordination, in short, "define the object of knowledge" (p. 56). Furthermore, he claimed, "[i]t seems obvious that these principles, which originate in reason, must be self-evident" (p. 57); in an endnote, he added: "it can easily be shown that the *insight into the necessary validity* of a priori propositions asserted by Kant does not differ from what we have called self-evidence" (p. 112).

In Reichenbach's view, several results follow from his proposed understanding of knowledge (or cognition) as coordination: A theory is considered to be true if, relative to the principles of coordination that it establishes, "all chains of reasoning lead to the same number [i.e., the same measurement] for the same phenomenon. This is the only criterion of [scientific] truth" (p. 43). In answer to the Kantian question, How is natural science possible, Reichenbach claims: "'Possible' is not meant in a psycho-physical, but in a logical sense: it pertains to *the logical conditions of a coordination*" (p. 47, italics added).

The relevance of Reichenbach's approach to the present study lies principally in his claim that the object as understood by science is a "reference structure": "[T]he conceptual scheme, the category, creates the object; the object of science is therefore not a 'thing-in-itself' but *a reference structure* based on intuition and constituted by categories" (p. 49, italics added). Physical reality—and, by extension (for the present author, and very likely also for Reichenbach) all objects of reference—can then be understood in what I shall later call a 'framework-relative' manner. As Reichenbach expressed this:

If the system of coordination is determined by reason in its conceptual relations, but in its ultimate construction by experience, then the totality expresses the nature of reason as well as the nature of reality; therefore, the concept of physical object is equally determined by reason and by the reality that the concept is intended to formulate. (p. 88)

Towards the end of his book, Reichenbach gave this summation:

We are offering this presentation of the concept of object of the theory of relativity—which makes no claim to exhaust the epistemological content of the theory—in order to show the

---

<sup>27</sup> Readers are reminded that here 'constitution' for Reichenbach has its Kantian meaning: an active, synthetic process that is the foundation of the very possibility of objective knowledge.



significance of constitutive [i.e., coordinative] principles. In contrast to particular laws, they do not say *what* is known in the individual case, but *how* knowledge is obtained; they define the knowable and say what knowledge means in its logical sense. Thus far they are the answer to the critical question: how is knowledge possible? (p. 104)

Reichenbach's book, as perhaps the reader is able to infer from the passages quoted, advocates, but does not demonstrate, an epistemological theory of the concept of physical object in terms of the way that concept has been transformed and re-defined by Einstein's theories of relativity: Both the special and the general theories assert the necessary and essential *linkage*, the relativity—of observations and measurements of objects and of physical states—to the physical conditions of any frame of reference in terms of which those observations and measurements are possible.<sup>28</sup>

In this work, I shall subsequently refer to this relativity as "*framework relativity*." The framework relativity of physical objects, which Reichenbach considered to be constitutive (in the transcendental sense) of those objects as "reference structures," foreshadows certain of the important results we shall reach as the metalogic of reference is developed.

Reichenbach, unfortunately in my view, situated his claims within a Kantian framework, a framework which, as is implicit in the passages quoted from his book, presumes that the knowing subject engages in an active, subjective constituting process. Kant's epistemological framework propounds that the object of scientific knowledge is *actively* constituted by categories—by coordinative principles—as required by the subject's reason. We shall find that such a claim is fundamentally unacceptable because it is referentially incoherent. In reaching that conclusion, certain of Reichenbach's informally expressed but central claims will take new and different forms, will be formulated in a more general and inclusive way, made more precise, and be provided with the justification that is absent in his book.

To review and re-state those principal claims:

- ◆ Physical reality can most effectively be conceived in coordinative terms
- ◆ Physical reality is defined by coordinative principles; they define what it is to be an object of knowledge

---

<sup>28</sup> A more detailed analysis is given later in {26}.

- ◆ Objects of knowledge therefore are not “things-in-themselves,” but structures that are determined by the referential framework presupposed in referring to them (i.e., they are “reference structures”)
- ◆ Coordination is the most general concept that describes the relation between concepts and reality
- ◆ Coordinative principles specify the preconditions that must be satisfied in order for knowledge to be possible
- ◆ Possibility in this sense has to do with the logical principles of coordination, which in some important sense are “self-evident”

...

After Reichenbach’s death in 1953, it was not until the 1960s that other philosophers directed attention to the theoretical importance of the concept of coordinate systems as applied to the theory of reference. The author’s doctoral dissertation, *A Relativistic Theory of Phenomenological Constitution: A Self-referential, Transcendental Approach to Conceptual Pathology*,<sup>29</sup> begun in 1965 and completed in 1970, remains the work that most extensively employs the concept of coordinate system in its fundamental role in theory of reference; mention of that study will occasionally be made. Later in the 1960s, Quine pointed, albeit only by inference and in passing, to the need to utilize the concept of coordinate systems in explicating reference:

[B]egin by picturing us at home in our language, with all its predicates and auxiliary devices.... This network of terms and predicates and auxiliary devices is, in relativity jargon, our frame of reference, or coordinate system. Relative to it we can and do talk meaningfully and distinctively.... [*R*]eference is nonsense except relative to a coordinate system. (Quine, 1969, p. 48, italics added)

Quine unfortunately did not develop this idea further—nor did he justify the categorical assertion made in the italicized last sentence, which nonetheless can be justified, as we shall see later.

---

<sup>29</sup> Bartlett (1970).

### 5.8 The metalogical study of reference: Preliminary comments

Having provisionally placed the naive view of reference to one side, we are in need of a replacement understanding of reference in order to proceed. In the last section, a basic claim about reference was made: that reference is essentially tied to identification; that any instance of reference is at once an instance of identification, whether that identification is vague or precise. The further claim was made that any object to which reference is possible *is* an object that possesses an identity—that to be a possible object of reference is to be an entity with some degree of identifiable identity.

These claims are validated when a coordinate point is specified, since the identity of that point requires an appropriate kind of coordinate system in order for the reference to obtain: The reference of a specified coordinate point is tied to its identifiability, while any specified coordinate automatically determines a specified point of reference. To be a specifiable coordinate point (a possible object of reference in an appropriate coordinate system) *is* to be a coordinate point with the specified identity.

In so describing the referential character of coordinate systems, there is, as the reader may have noted, a sense in which some variety of what one might call ‘tautological circularity’ has begun to come into view: A description of the logically interdependent relations among coordinate system, specification of coordinate point, and coordinate point identity is descriptive of coordinative relations that exhibit the character of an interconnected, dynamic, general system.

This intuitive sense is close to the mark. Considered as a “reference structure,” to use the phrase adopted by Reichenbach, an object possesses this “closed circuit” variety of logical dynamic. In much the same way that the very identities (i.e., the measurable physical properties and relations) of physical objects and events constitute possible objects of knowledge from the standpoint of relativity physics, so do objects of reference, considered generally, constitute possible objects of knowledge from the standpoint of the coordinate systems required to identify them. In Reichenbach’s terms, the coordinative principles that make reference possible to physical objects establish a functional relationship between coordinate system and the measurable properties of physical objects that can, in principle, be observed from the standpoint of that system. This is an interdependent, inalienable relationship: The identity of such physical objects is inseparable, and logically meaningless as Quine suggested, apart from the coordinate system or systems that permit reference to them, as we shall later show in detail.

It should immediately be clear how, from this abstract level of analysis, the object of reference and its specified identity are not disjunct, but are

essentially fused: There is no possible object of reference without a specific identity (whether that identity is vaguely or precisely specified).

Occasionally in the literature one finds some passing recognition of this metalogical observation. Quine proposed a concise dictum: “no entity without identity” (Linsky, 1967, p. 27). His maxim recalls Wittgenstein’s idea that “[i]t is impossible to represent in language anything that ‘contradicts logic’ as it is in geometry ... to give coordinates of a point that does not exist” (Wittgenstein, 1961/1921, §3.032).<sup>30</sup> Any entity whatsoever has some sort of identity, else reference to it could not obtain. The two—possession of identity and the possibility of reference—are intimately linked and inseparable.

Statements such as these have seemed, for good reason, to be reliable truths, but they are seldom developed and even less seldom, if ever, proved.

What I call ‘metalogical reference’ is reference considered on this abstract and general level. Unlike pragmatism reference described earlier in this chapter, a study of metalogical reference seeks to make explicit the referential preconditions that must obtain in order for reference to be possible within a particular frame of reference, or within a class of such reference frames.

---

<sup>30</sup> Physicist and philosopher P. W. Bridgman made an interesting, related claim: “In our dealing with the world of objects there is one operation which we always assume can be performed, namely, the operation by which we attach an identity to the object. Some such concept seems to be almost a ‘necessity of thought’ in dealing with the world” (Bridgman, 1959, p. 44). —This claim, however, though close to the mark as were Quine’s and Wittgenstein’s, falls short and is misleading: To suggest that an activity occurs that “*attaches* an identity to the object” is to make a referentially incoherent claim, as is later shown. Here, again, we encounter what I regard as the philosophical pitfall of equating reference with some kind of performative “operation.”

## 6

# Self-referential Argumentation and the Metalogic of Reference

### 6.1 Self-referential argumentation in philosophy

Within philosophy, the phenomenon of self-reference or reflexivity has inspired research in three main areas: in semantic theory, theory of argument, and theory of knowledge. Of these, the earliest studies of reflexivity were made in semantic theory. They sought to understand the impact of the paradoxes encountered in number theory and the theory of classes upon the capacity of propositions, both those of formal systems and those in non-formalized discourse, to assert truth without self-referential inconsistency.

Shortly thereafter, a small group of philosophers began to cultivate an explicit interest in the use of self-reference in philosophical argument. Although individual examples of reflexive argumentation have peppered the history of philosophy, it was not until the middle of the twentieth century that efforts were made to construct a theory of self-referential argumentation. This work has largely been carried out by philosophers whose principal focus in argumentation has been the use of language. Several varieties of language-based self-reference relevant to philosophical argumentation have been identified. Of these, two main divisions may be distinguished: (i) contingent forms of self-referential statement that are tied to factual conditions surrounding a speaker's assertion, and (ii) self-referential propositions that do not depend upon the factual conditions of their use.

Interest in the former "pragmatic"<sup>31</sup> or "performative" variety has dominated the literature, studying, for example, self-referential statements whose content, if in fact asserted, conflicts with the performative conditions of that assertion, thereby falsifying it (e.g., *P*'s assertion, "I can't assert anything"); similarly, self-referential statements whose content, if in fact asserted, confirms the truth of the assertion (e.g., *P*'s assertion, "I exist"); semantically self-referential statements (such as *P*'s statement "All statements made by *P*

---

<sup>31</sup> Alternatively called 'pragmatic' by some authors.

are true”); etc.

“Propositional self-reference” may be distinguished from the pragmatical variety; we find propositional self-reference in propositions that refer to themselves no matter how they are stated (e.g., the propositions, “All propositions are either true or false,” “It can be proved that nothing is provable,” “All propositions are false”). As has been mentioned, less attention has been devoted to this kind of self-reference. It is, nonetheless, an important variety of self-reference, one that is involved whenever a position itself—considered independently of the factual conditions of utterances or exchanges between speakers and hearers—is shown, in and by its own terms, to be self-referentially inconsistent.<sup>32</sup> Proofs by means of propositional self-reference can therefore be characterized as “activity-independent,” in contrast to pragmatical arguments.

A wide range of distinctions has been made among kinds of self-referring statements and propositions; arguments have been advanced concerning the legitimacy or illegitimacy of certain of these; and numerous proposals have been advanced that recommend how these different kinds of self-reference may be employed in philosophical argumentation. The reader who is interested in the multitude of kinds of self-reference is referred elsewhere.<sup>33</sup>

A third area of philosophical interest inspired by reflexivity has evolved from the Kantian and Husserlian attempts to identify the transcendental preconditions of objective knowledge. Here, the internal limitations of human self-understanding become especially evident in the human effort to acquire knowledge about the limits of what human subjects can know. In previous publications, I have called the variety of reflexivity that is relevant to this task ‘*metalogical self-reference*’. Studies of metalogical self-reference describe the general and necessary conditions that underlie the capacity of our concepts, of our theories and claims to knowledge, in principle, to refer at all, no matter what the object may be to which there may be reference.

As the earlier chapters in this book have observed, progress in philosophy has been impeded by the absence of a shared methodology among philosophers, by excessive tolerance among philosophers for statements of mere belief, and by a psychology of intellectual recalcitrance and narcissism among philosophers. Without an impartial, objective set of standards as is

---

<sup>32</sup> Among those who have recognized this variety are, e.g., Mackie (1964, pp. 195ff) and Boyle, Grisez, and Tollefsen (1972, p. 20): “[A] position can be shown to be self-referentially inconsistent even if no one actually asserts it.”

<sup>33</sup> See Bartlett (1987) for an overview of varieties of self-reference. For an extensive bibliography of the literature, divided according to the types of self-reference studied, see Suber (1987). Boyle, Tollefsen, and Grisez (1976) devoted Chapter 5 to a basic description of kinds of self-referential statements and propositions, and ways in which these can go wrong.

presupposed in science and mathematics, without disciplinary constraints that insist upon the exclusion of expressions of mere belief, and dominated by a psychology that fosters contention among competing views, philosophers have relied primarily upon two approaches to philosophical argumentation: most commonly, appeal to logical standards of validity, and, less often, application of self-referential argumentation.

The ability to appeal successfully to the standards of valid argument established by traditional logic—that is, the ability to justify or to criticize a given argument using logical consistency as an arbiter—has been very substantially weakened by realizations reached largely during the twentieth century that even the most basic laws and rules of logical inference are themselves open to question and that alternatives to them have become available. Divergent systems of logic have multiplied, including those that are many-valued, or are inconsistency-tolerant (when less informative, stylish names are preferred, they are called ‘paraconsistent’), or are inconsistency-affirming (AKA ‘dialetheic’). This multiplication of systems of logic has widened the range of often incompatible standards of logical validity that philosophers may choose among when seeking to argue either on behalf of their beliefs and claims to knowledge, or against views they wish to defeat.

Earlier in this book, I described from a psychological perspective how a profession’s self-selecting psychology of intellectual self-absorption and promotion of individual beliefs has led to a discipline that cannot help but be dominated by contention and competition among conflicting views, pursued by members of the profession the majority of whom have a vested interest in urging that their own ideas be given attention and priority, to the exclusion of others. Attention and popular priority have become more important, in part because they are clearly more attainable than demonstrated truth.

Because it possesses no objective, externally established tools by means of which to convince fellow philosophers of the errors of their views and the truth of one’s own, such a discipline is forced to rely very heavily upon approaches to argumentation that take on an opponent philosopher’s view *in its own terms*, that is, *from within* the opponent’s framework.

Recourse to internally established standards is an old method of argumentation, found already in use by the ancients, but in the last century it has led to the development of an explicit metatheory of philosophical argumentation. One of the most dedicated metaphilosophers with this interest was Henry W. Johnstone, Jr. (1920-2000), who sought to show that all valid philosophical arguments are basically *ad hominem*—not in the ordinary sense of arguments that rely on personal attacks upon an opponent, but in the sense in view here of argumentation that seeks to situate itself within an opponent’s framework.

*Ad hominem* argument, in this meaning, attempts to show, in any of a variety of ways, that an opposing philosopher's stated position defeats its own purposes.

There are numerous ways in which this can be attempted—for example by showing that a given argument reaches conclusions incompatible with its own presupposed principles of reasoning or incompatible with its own asserted content; the argument may be criticized for internal ambiguity or inconsistency, or for failing to present evidence adequately or fairly, or simply for failing to show what it claims it intended to show; the argument can be faulted for presupposing what its own assertions directly or indirectly deny; or it might be claimed that the argument does not, by its own standards, reach results that can qualify as truths or knowledge; or the argument may be rejected on grounds that it undermines its own truth; or that it presupposes the truth of its own conclusion; etc.<sup>34</sup>

If one grants that, in general, philosophers characteristically are monadic in their philosophical studies, that the views they come to espouse reflect this same property of monadic self-enclosure, i.e., of resistance to criticism formulated from outside their preferred frameworks, then the use of *ad hominem* argumentation in Johnstone's sense is not only understandable, but essential. Within such a disciplinary domain consisting largely of self-encapsulated views, the only acceptable evaluative means will be one that functions through internal criticism.

There are two main approaches to internal philosophical criticism: One is *rhetorical* in the classical meaning of this word: its purpose is to *persuade* others to accept the validity of one's arguments and the truth of the conclusions to which they lead. The rhetorical approach is therefore *person-directed*: Arguments are targeted at individuals or a certain audience. Rhetorical argumentation is something of an art form, requiring a skillful apprehension of an opponent's strengths and weaknesses, keenly honed abilities both to make often small-scale adjustments in one's modes of expression and to engage in on-the-toes repartee, intellectual and psychological agility in adjusting to the changing contexts of interpersonal exchanges, etc. Rhetorical argumentation is most fundamentally *performative*, and hence applications of self-referential argumentation that tend to lend themselves most readily to rhetorical use are pragmatical forms of self-referential argumentation. We shall look more closely at these in a moment.

The other main approach to internal philosophical criticism is *conceptual* rather than rhetorical; its purpose is disassociated from the give and take, the thrust and counter-thrust, of interpersonal exchanges, and instead develops

---

<sup>34</sup> See, e.g., Johnstone (1970), which contains references to others of his publications.



arguments that are indifferent to and independent of an opponent's or a group's allegiances. Mathematical argument is of this kind: It is characterized by explicitly understood and accepted meta-rules—for example, that it is considered “beside the point” to question whether the developer of a particular formal system has a right to stipulate certain axioms that are to be taken for granted. Instead, evaluation of that formal system is a matter of judging whether the reasoning employed and the results reached are valid based on the rules of inference which that system permits. The allegiances of other mathematicians, if they have allegiances that lead them in their own work to favor alternative approaches, are not *mathematically* relevant (although they often exert considerable psychological force, even among mathematicians). Similarly, the evaluation of theories in natural science is, ideally, disengaged from the personal adherences of individual scientists, who must bow, often with great reluctance, to the weight of accumulating empirical evidence.

Philosophical argumentation that is conceptually rather than rhetorically focused is similar to the mathematical approach. Arguments that are formulated with conceptual evaluation in view, in marked contrast to arguments intended to respond in a dynamic fashion to exchanges between persons, may be likened to the hypothetical model of reasoning since they establish their starting points in a stipulative manner. Doing this removes what might be called the ‘interpretation liability’ incurred whenever argument exchanges between persons occur; it avoids the vagueness and shifting grounds of rhetorical argumentation; and it permits a more rigorous approach. A conceptual argument will typically have this form: *Given* a designated starting point, which stipulates the subject-matter and framework at issue, and then applying rules of inference that are endorsed, certain conclusions necessarily follow. As we shall explore in greater detail later, this approach is fundamentally what I shall call ‘*framework-relative*’, a term I shall use in preference to ‘hypothetical’, which can erroneously suggest that unreliable or provisional claims are involved. It is an approach that is essentially *respectful* of the framework to be analyzed; the results reached are, as I have previously expressed this, “*intimately linked*” to the frame of reference that is in view.

We note, then, a significant difference between rhetorical philosophical argumentation and the conceptual variety: Making a rhetorical philosophical argument “work” to show that an opponent's position is mistaken or unacceptable is like trying to hit a moving target. Johnstone's observation that “man has an unlimited capacity for sophistry” comes to mind. “To put the point more directly, there is no argument, valid or not, that can altogether silence a philosopher if he wishes to continue the discussion. However devastating the attack upon him, there is always something more he can say”

(Johnstone, 1964, pp. 480-481). A conceptually focused philosophical argument, in contrast, aims at a fixed and stable target, one which does not morph into something different when criticism is brought to bear on it. Where success in the first instance is nearly always questionable and open to further debate, in the second it is not.

The variety of “linkage”—between the results reached and the frame of reference that is in view—which will concern us in what follows has, then, to do with maximally general, theoretically abstract preconditions of reference that must be granted if reference is to be possible relative to a given frame of reference. The approach to conceptual analysis developed in this study will therefore make fundamental use of the metalogical variety of self-referential argument.

## 6.2 Pragmatical, or performative, self-referential arguments

When someone makes an assertion, there are two dimensions or aspects of the assertion that may or may not conflict with one another. One aspect concerns the *content* of the statement, what it asserts; the other has to do with the *factual way* in which the statement is asserted, that is, with *how* the speaker intends the statement to be understood by the person(s) to whom the statement is addressed. A statement that is asserted in such a way that these two dimensions refer one to the other is typically called ‘*pragmatically, or performatively, self-referential*’. For example, if a speaker asserts the truth of the claim, “There are absolutely no truths at all,” that statement is self-referentially inconsistent in the pragmatical or performative sense. If I say, with perfect enunciation, “I just can’t say the word ‘antidisestablishmentarianism,’” or say aloud, “I am unable to use my voice,” or assert with obvious intended meaning, “I am unable to say anything meaningful,” I’ve said something that is pragmatically self-refuting.

Pragmatical self-reference directs attention to the *factual commitments or conditions* involved in making an assertion. To take another example, the assertion, “Knowledge is impossible in this world of flux,” intended by the speaker to be regarded as itself a knowledge claim, is pragmatically self-referentially inconsistent: Provided that the assertion is *in fact* linked to this underlying commitment on the part of the speaker, a commitment that places the assertion in the category of knowledge claims, the assertion is pragmatically self-refuting. The challenging task of the pragmatical self-referential analyst is to reveal the existence of the factual commitments that underlie everyday and philosophical discourse. His results stand or fall depending on the convincingness of his factually focused demonstration.

Pragmatical self-referential arguments have tended for the most part to be critical, negative, or corrective, seeking to identify and eliminate internally inconsistent claims. Some pragmatical applications, however, have sought to use self-referential approaches to establish results non-destructively. In what follows, we look at both kinds of applications.

### 6.3 The critical use of pragmatical self-reference

The strange thing is that philosophers should have been able to hold sincerely, as part of their philosophical creed, propositions inconsistent with what they themselves knew to be true; and yet, as far as I can make out, this has really happened.

– G. E. Moore (1965, pp. 53-54)

Pragmatical applications of self-reference have attempted to show that such claims as these are self-falsifying:

- ◆ Pleasure is the chief good, since any good thing is made more desirable by the addition of pleasure.<sup>35</sup>
- ◆ The materialist can explain the causes of our ideas in terms of external bodies.<sup>36</sup>
- ◆ Every event must have a cause.<sup>37</sup>
- ◆ All knowledge, including this, is a product of an organism's adjustment to its environment.<sup>38</sup>
- ◆ All meaningful statements are verifiable.<sup>39</sup>
- ◆ Science is incapable of objectivity.<sup>40</sup>
- ◆ The shift from one theory to another involves an incommensurable change in the meanings of the terms used, so that there cannot be any statements invariant across theories.<sup>41</sup>

---

<sup>35</sup> Argument from Eudoxus; see treatment by Johnstone (1970, pp. 64ff).

<sup>36</sup> Johnstone (1970, pp. 67ff).

<sup>37</sup> Argument from Hume; see discussion in Johnstone (1970, p. 95).

<sup>38</sup> Urban (1949, pp. 69ff).

<sup>39</sup> Rorty (1961, esp. pp. 104-107).

<sup>40</sup> Kordig (1970).

<sup>41</sup> Kordig (1970).

- ◆ No hypothesis can be immune to revision.<sup>42</sup>
- ◆ No hypothesis can be irrevocably falsified.<sup>43</sup>
- ◆ All our statements lack significance.<sup>44</sup>

To this short list could be added many other examples, for numerous philosophical positions have been indicted for falling victims to the pragmatical variety of self-referential inconsistency. Among those that have been attacked in this way are the coherence theory of truth,<sup>45</sup> pragmatism,<sup>46</sup> skepticism,<sup>47</sup> intuitionism,<sup>48</sup> behaviorism,<sup>49</sup> determinism,<sup>50</sup> subjectivism,<sup>51</sup> views that oppose idealism,<sup>52</sup> and views that oppose utilitarianism.<sup>53</sup>

#### 6.4 The constructive use of pragmatical self-reference

[V]alid constructive arguments in philosophy must in fact be circular.... All valid constructive philosophical arguments involve this element of feedback.

– Henry W. Johnstone, Jr. (1970, pp. 76, 78)

In contrast to the preceding variety of critical arguments that utilize pragmatical self-reference to undercut an objectionable thesis, a few philosophers have tried to use the approach constructively.

(We should remark that this distinction, between critical and constructive arguments, is often difficult to draw clearly, especially in the present context: A pragmatically critical argument establishing that *P* is self-falsifying may be judged to lead to the conclusion *not-P*; yet, if *not-P* is thought to be a philosophically significant result, the argument's proponent naturally believes his

---

<sup>42</sup> Kordig (1970).

<sup>43</sup> Kordig (1970).

<sup>44</sup> Passmore (1961, p. 69).

<sup>45</sup> Spaulding (1918, pp. 350-351).

<sup>46</sup> Royce (1904, pp. 128-129).

<sup>47</sup> Urban (1929, pp. 45-46) and Passmore (1961, pp. 72ff).

<sup>48</sup> Hocking (1939, p. 201).

<sup>49</sup> Lovejoy (1922, pp. 142-147).

<sup>50</sup> Lucas (1970) and Boyle, Grisez, & Tollefsen (1976).

<sup>51</sup> An argument originally advanced by Protagoras: see treatment in Passmore (1961, pp. 64ff).

<sup>52</sup> Royce (1919, pp. 237-240; 1959/1899, pp. 136-138).

<sup>53</sup> Bentham (1876, Chap. 1, sections 13-14).

argument is constructive. Among arguments and their proponents, the constructiveness of their conclusions can be stretched across a broad spectrum. At the dim end of lesser interest one might place, for instance, the critical argument against the assertion, "All our assertions lack significance." For many people, a self-referential argument that claims to establish the negation of this assertion does not seem especially interesting or constructive. Certainly it tells us something of which few are ignorant.)

In general, constructive self-referential argumentation attempts to demonstrate a positive thesis, rather than to undermine an erroneous view maintained by someone else. Admittedly, the judo-like strategy of utilizing feedback in argumentation is especially well-suited to showing that a position that an opponent holds is wrong, as a reader new to the field intuitively may suspect. But some constructive arguments have, nevertheless, been formulated using the tools of pragmatism and self-reference. A few we might mention here are:

- ◆ Moore's defense of common sense, using its appeal;<sup>54</sup>
- ◆ The argument that there are invariant conditions of discourse;<sup>55</sup>
- ◆ Arguments seeking to demonstrate the ontological commitments of discourse, and the related argument claiming that all objects of which we are conscious are, in diverse senses, real;<sup>56</sup>
- ◆ The self-confirming evidence that a sound is audible, is that we hear it;<sup>57</sup>
- ◆ The defense of "orientational pluralism" in philosophy: According to this view, philosophical positions represent relativist frames of reference. For them, there is no unique solution to philosophical problems.<sup>58</sup>

To these examples may be added the larger group of arguments that progress

---

<sup>54</sup> Passmore (1961, pp. 78ff).

<sup>55</sup> Passmore (1961, pp. 69ff); Lorenzen (1969, p. 14; and in connection with operative logic, p. 89). See also Lorenzen (1969a).

<sup>56</sup> Quine (1969) and Meinong (1960/1904).

<sup>57</sup> Mentioned by Mill (1971/1863, Chapter IV); also see discussion in Johnstone (1970, pp. 77ff).

<sup>58</sup> Rescher (1979, pp. 217-251).

from a self-referential refutation of an opposing thesis to the affirmation of its philosophically significant negation. Among these are found the positions mentioned earlier that defend: the objectivity of science, free choice, utilitarianism, idealism, the thesis that verifiability is not a property belonging to all meaningful statements, etc.

Representing something of a transition between the category of pragmatiscal self-referential arguments and that of metalogical self-referential arguments are the approaches to self-referential argumentation proposed separately by John Passmore, J. L. Mackie, and Gaston Isaye. We take a few moments to look at their work in detail.

### 6.5 Passmore's absolute self-refutation

My purpose here is not to provide a detailed exposition and criticism of Passmore's contribution to the topic of self-referential argumentation, but rather to highlight briefly where his thought in this area of study may most appropriately be placed in the gamut from pragmatiscal to metalogical argumentation.

In his 1961 book, *Philosophical Reasoning*, Australian philosopher John Passmore (1914–2004) distinguished three varieties of self-refutation that philosophers may use in evaluating philosophical positions:

- ◆ *Pragmatic self-refutations* involve a conflict between a statement and certain facts surrounding its statement. For example, when spoken by someone, “I cannot speak” is pragmatically self-refuting.
- ◆ *Ad hominem self-refutations* involve a conflict between a statement and certain admissions made by the person whose statement it is. When person X claims “I cannot speak” while admitting that he *is* speaking, his statement comprises an *ad hominem* self-refutation.
- ◆ An *absolute self-refutation* occurs, in Passmore's view, when what a statement asserts conflicts with presuppositions made by all assertions—“the presuppositions of all proposing” (Passmore, 1961, p. 68), which he called ‘invariant conditions of discourse’ (p. 80).<sup>59</sup> We shall discuss this variety in more detail in what follows.

---

<sup>59</sup> Passmore's concept of absolute self-refutation very likely had its roots in the thought of his

Passmore claimed that pragmatic self-refutations can always be evaded: The propounder of a position under criticism “can always—in principle, even if sometimes with almost inconceivable hardihood—deny that he has in fact pronounced the word or uttered the statement [in question]” (p. 63). But with absolute self-refutations, it is another matter: “only if a philosophical argument can show ... that a sentence can propose nothing—because *what it asserts, if it were taken to propose something, would be inconsistent with the presuppositions of all proposing*—is it pointing, I suggest, to an absolute self-refutation” (68, italics added). He claimed: “A view is absolutely self-refuting only if it is incompatible with [the] invariant conditions of discourse” (p. 80).<sup>60</sup>

One of the examples Passmore gave is the universal proposition “All universal propositions are false,” which in his view refutes itself absolutely because, if it were true, it would contradict itself, but, more importantly as I understand his intent, because it reveals, in his view, an incompatibility with a condition of discourse that cannot be surrendered without crippling, and indeed making discourse itself impossible. “Absolute self-refutation [according to Passmore] depends not upon the fact that the defendant is thinking or speaking, or that he is admitting something, but upon *the fact that he is making an assertion that denies the very possibility of making an assertion*” (Johnstone, 1964, p. 471, italics added).

Another example Passmore gave is “No sentence conveys anything,” about which he wrote:

---

teacher, John Anderson, a philosophy professor at the University of Sydney. In 1927, Anderson wrote, “The chief, and I think final, objection to any theory of higher and lower, or complete and incomplete, truth is that it is contrary to the very nature and possibility of discourse; that it is “unspeakable” (Anderson, 1962, p. 4). —Anderson used the term ‘unspeakable’ as synonymous with ‘self-refuting’ (p. 12). In 1936, he recommended an approach to philosophical argumentation “by considering what is involved in the recognition of a thing as a subject of investigation—more generally, in the very possibility of ‘discourse’.... [I]n rejecting a particular logical theory, we should be able to show that the exponent of it not merely has a false view of existence but implicitly, in his own statement of the case, admits the view that we are upholding against him (as when a person *argues* against objective implication or *denies* objective truth).... [I]ndirect ‘proof’ of a logical position may take the form of showing that our opponent’s view involves him in insoluble problems—though this amounts to the same as contradicting the possibility of discourse” (p. 123).

<sup>60</sup> At other times in the same book, however, Passmore proposed a decidedly different, and to my mind unfortunate, formulation, one that took him in a direction away from the presuppositional-invariant purpose of absolute self-refutations: In this second formulation, a proposition *p* is absolutely self-refuting “if to assert *p* is *equivalent* to the assertion *both p and not-p*” (60, former italics added). It is one thing for a proposition simply to entail a logical contradiction, and another for that proposition to conflict with “the presuppositions of all proposing,” for certainly not all propositions that entail contradictions are propositions of the latter sort.

[W]e cannot renounce the claim that some of our statements are significant.... To invite us to discuss ‘no sentence conveys anything’ is to ask us to consider whether *what that sentence conveys* is true. To assert that no sentence conveys anything is absolutely self-refuting: for “‘no sentence conveys anything’ is true” asserts that “what the sentence ‘no sentence conveys anything’ conveys is true, and ‘no sentence conveys anything’ is true” (69-70).<sup>61</sup>

Passmore observed that “to engage in discourse at all [one] has to assert that something is the case” (67), and so “it is presupposed in all discourse that some propositions are true, that there is a difference between being the case and not being the case, and to deny this in discourse is already to presume the existence of the difference—since otherwise the notion of ‘denying’ is quite meaningless” (68).

Passmore unfortunately did not provide a more fully developed account of absolute self-refutation, and, as a result, has been criticized, sometimes without a wholly sympathetic understanding of the direction he was taking.<sup>62</sup> Certainly his account, as I see it, falls short because he did not explain clearly and fully what “the presuppositions of all proposing” or “invariant conditions of discourse” really are, how they can be identified, and how they can be demonstrated.

It is tempting to subsume his concept of absolute self-refutation under the category of pragmatical self-referential argumentation,<sup>63</sup> but even the short passages quoted above do not seem to me to support this. He rather seems to have been on the path toward recognizing certain general and invariant principles that can be appealed to in order to justify philosophical arguments that involve self-refutations, i.e., those kinds of arguments that *cannot* be evaded with the “almost inconceivable hardihood” that afflicts pragmatical self-refutations. If this interpretation of his intentions is valid, then placing his work in the “transition area” between pragmatical and metalogical self-referential argumentation is appropriate.

---

<sup>61</sup> Here, as he often did, Passmore identified absolute self-refutations with propositions that are logically equivalent to contradictions, and in doing this he lost sight of what I think was his main interest—invariant presuppositions of all discourse. See the previous note.

<sup>62</sup> See, for example, Johnstone (1964).

<sup>63</sup> This is essentially what Johnstone did: see Johnstone (1964, pp. 476f).



### 6.6 J. L. Mackie's operational self-refutation

Like Passmore, J. L. Mackie (1917–1981) was an Australian philosopher, and like Passmore was strongly influenced by their teacher, philosopher John Anderson.<sup>64</sup> Mackie formulated what he claimed is a distinct variety of self-refutation, “operational self-refutation.” He presented this idea in a worthwhile but not often or adequately appreciated paper which I would like to take note of here.<sup>65</sup>

This paper (Mackie, 1964) is of special interest because it seeks to provide a rare attempt to formalize distinct forms of self-refutation. To do this, Mackie made the following definitions:<sup>66</sup>

- $d$  is a proposition-forming operator (standing, e.g., for “I assert that” or “it can be proved that”);
- $d$  is called ‘*truth-entailing*’ when  $dp \rightarrow p$  (e.g., when  $d$  is: “It can be proved that,” “I know that,” “It is true that”);
- $d$  is called ‘*strictly prefixable*’ when  $p \rightarrow dp$  (e.g., when  $d$  is: “It is possible that” or “It is true that”; the latter is also truth-entailing, but not the former);
- $d$  is called a ‘*weakly prefixable*’ operator when there is an implicit commitment to some claim or condition (e.g., the assertion of  $p$  is an implicit commitment that  $p$  is known, so that “I know nothing” is an implicit commitment that “I know that I know nothing”); and
- Let  $a$  symbolize “ $x$  coherently asserts that,” which is to mean that  $x$  cannot coherently assert a self-contradiction, and that if  $p$  entails  $q$ , then  $ap \rightarrow aq$ . If  $x$  coherently asserts that  $p$ , and if the coherent assertion of  $p$  implicitly commits  $x$  to asserting  $q$ , then it follows that  $x$  coherently asserts  $q$ .

---

<sup>64</sup> See note 58.

<sup>65</sup> One of the few to recognize the importance of Mackie’s contribution is Castagnoli (2007): “More than four decades after its first publication in 1964, Mackie’s formal analysis of the logic of self-refutation remains the best ... on account both of its undeniable merits and of the scarcity of subsequent attempts” (p. 13).

<sup>66</sup> In his paper, Mackie used Polish outfix notation, which I have here translated into the more commonly employed infix notation.

With these preliminaries, Mackie identified four varieties of self-refutation:

(1) *pragmatic refutation*:  $d \sim \exists pdp$ , when this symbolizes the *factual situation* in which I write that I am not writing, a claim that is self-falsifying;

(2) *truth-entailing absolute self-refutation*: when the *proposition* (and not the factual situation of performance) expressed by  $d \sim \exists pdp$  is absolutely self-refuting—e.g., “It can be proved that nothing can be proved” must be false;

(3) *strictly prefixable absolute self-refutation*: when the *proposition* symbolized by  $\sim \exists pdp$  is absolutely self-refuting—e.g., “There are no truths”; and

(4) *weakly prefixable operational self-refutation*: when the *proposition* expressed by  $\sim \exists pdp$  cannot coherently, i.e., without self-contradiction, be asserted—e.g., I cannot coherently assert “I believe nothing” or “I know nothing.”

Operational self-refutation, in Mackie’s view, is then intermediate in strength between pragmatic and absolute self-refutation:

In pragmatic self-refutation the way in which an item happens to be presented conflicts with the item itself. But where we find operational self-refutation *there is no other way* in which the precise item can be presented. The only possible way of presenting the item is to “coherently assert” it, and since this involves asserting something that conflicts with the item itself, this precise item cannot be presented at all. (Mackie, 1964, p. 197, italics added)

On the surface, Mackie’s distinction between pragmatic (what I’ve called ‘pragmatical’) and operational self-refutation may seem to cut things too finely, but this is not the case. Where a pragmatic self-refutation focuses on the factual conditions surrounding a particular linguistic performance, the operational variety is intended to relate to the “implicit commitments” of coher-

ent assertions,<sup>67</sup> and in this sense, might suggest a constructive step beyond Passmore's loosely defined "invariant conditions of discourse," a step that was not, as will be made clear, as fully developed by Mackie as it might be.

Mackie proceeded in his paper to discuss the applicability of these different forms of self-refutation in the context of brief analyses that include Descartes' *cogito*, skepticism, and Berkeley's self-refutation argument concerning the independent existence of material things (I will not discuss these analyses here as this would take us too far afield).

In Mackie's view, although there are indefinitely numerous possible instances of pragmatic self-refutation, he judged them to be of limited philosophical value:

Nothing much follows from such facts as that if I write that I do not write, then what I write is false, and even in more complicated cases the most that is shown is that a certain way of presenting a certain view is unsatisfactory, that a proposed supporting argument will not support the view it is intended to support; but the view itself is not thereby refuted, and it may well be presented and supported in other ways. (Mackie, 1964, pp. 202-203)

Similarly, he noted that although there are many truth-entailing operators, their application does not tend to yield important results: "[They] tell us only such things as 'It cannot be proved that nothing can be proved' " (p. 203), a conclusion that isn't particularly informative. With respect to strictly prefixable operators, he claimed that there are a very limited number of absolutely self-refuting propositions of the form  $\sim\exists pdp$ , such as "It is not the case that something is possible" (p. 195).

If these observations are true, then, Mackie argued, the principal, philosophically significant use of self-refutation is to identify propositions that are operationally self-refuting, propositions that cannot coherently be asserted

---

<sup>67</sup> Although in his paper Mackie employed the word 'presupposition' in discussing the work of G. E. Moore and John Passmore, he seemed carefully to avoid using this term when explaining his concept of the "implicit commitments" of coherent assertions. Perhaps he did this to distance himself, for example, from Passmore's "presuppositions of proposing" which Passmore had characterized as "the invariant conditions of discourse" (Passmore, 1961, pp. 68 and 80, referred to in Mackie, 1964, p. 202). And yet Mackie's "implicit commitments" clearly are intended to refer to "presuppositions" that are implicitly implicated by coherent assertions involving weakly prefixable operators. In other words, "implicit commitments" are intended to refer to presuppositions that are unavoidably involved when, e.g., claims of belief and knowledge are made.

because they conflict with “implicit commitments,” commitments that simply cannot be avoided when such assertions are coherently made. However, he was explicitly modest in his expectations that such a task would lead to philosophically noteworthy positive, necessary truths.

The fact, for example, that “I believe nothing” cannot coherently be asserted does not lead, beyond this result, to any important consequence. Nor, in Mackie’s analysis, does the application of operational self-refutation enable us to establish as logically necessary truths Descartes’ proposition that “I am (essentially) a thinking being,” or Berkeley’s claim that “Material objects do not exist unconceived.” Operational self-refutation cannot provide the justification Descartes and Berkeley wished; at most, propositions that deny these central claims by Descartes and Berkeley are not absolutely, but only operationally, self-refuting. Mackie concluded “the detection of [these operationally self-refuting propositions] does not lead to such necessary truths [as Descartes and Berkeley claimed]” (p. 203).

Referring to the often-sensed, apparent logical force behind Descartes’ *cogito* and behind Berkeley’s claim that to assert the independent existence of material objects is to assert a contradiction, Mackie concluded: “On the contrary, we are now in a better position to understand why these propositions have seemed to be necessary and to criticize the arguments by which philosophers have tried to establish them” (p. 203). In his view, these propositions have seemed to be necessary “only” because denying them leads to what he identified as operational self-refutations.

In the end, for Mackie, self-refutation, as it is expressed in all four varieties he identified, remains *only* a tool of negative criticism, not a promising means of establishing significant philosophically positive, necessary results. Here is the way he summarized this, referring to Passmore’s “invariant conditions of discourse” (note his explicit reservations, expressed in the twice-repeated word ‘only’):

There may indeed be conditions of discourse, but they seem to be of more than one type. The doctrine that all truths are relative is absolutely self-refuting [in Mackie’s sense (3), earlier], because the condition of discourse which it violates is the fact that ‘It is an absolute truth that’ is a strictly prefixable operator. But in general what could be called a condition of discourse *sets limits only to what can be coherently asserted*; it is concerned with the way in which saying one thing implicitly commits us to being prepared to say something else as well, and it therefore gives rise *only* to what we have called

operational self-refutation. (p. 202, italics and underscoring added)

This is where Mackie left his concept of operational self-refutation. But notwithstanding his modest claims on its behalf, in my estimation the importance of Mackie's operational self-refutation lies in the general observation that in ordinary discourse and especially in philosophical thought commitments are implicitly in force, commitments upon which the very *coherence* of such discourse or thought depends. The "limit-setting" dependency relation that would be in question here—were one to develop further the analysis provided by Mackie, and which in a way we shall try to do in this study—is a relation that underlies all possible coherent assertions.

As Mackie wrote in the passage just quoted (I have removed one 'only'): "what could be called a condition of discourse *sets limits ... to what can be coherently asserted.*" As we shall find later on, there is a good deal more to such "limit-setting" forms of presuppositional relation than Mackie recognized.

### 6.7 Isaye's *rétorsion*

As in the foregoing brief discussions of Passmore and Mackie, my comments about Isaye's work are intended primarily to situate his thought in terms of the range of approaches to philosophical self-referential argumentation.

Gaston Isaye (1903-1984) was a Belgian philosopher and Jesuit priest whose publications in epistemology are less known to Anglo-American readers, in large part due to their lack of translations into English.<sup>68</sup> Isaye's method of self-referential argumentation was deeply influenced by the thought of Joseph Maréchal, who was responsible for introducing Kant's transcendental method into Thomistic philosophy.<sup>69</sup>

Isaye, like Maréchal, considered the role of affirmation to be central and indeed unavoidable to epistemology. Both philosophers observed that no claims to knowledge, nor claims by epistemologists about such claims to knowledge, are possible without the necessary involvement of affirmation. Even efforts to avoid affirming must themselves necessarily comprise affirmations. As Joseph Donceel expressed this in his commentary on Maréchal's thought: "... *the supreme effort of the human mind to keep away from affirm-*

---

<sup>68</sup> Isaye's philosophical work has very likely also suffered from neglect by mainstream philosophy due to the often justifiable prejudice against Catholic Thomists, in whose tradition Isaye is normally placed, who use philosophy as a handmaiden for religious dogma.

<sup>69</sup> See Joseph Donceel's Introduction to Maréchal (1970), also Donceel (1974).

ing is one more affirmation. Hence affirmation is unavoidable” (Maréchal, 1970, p. 9, original italics).

The basis for Isaye’s self-referential method dates back to Aristotle and St. Thomas. Maréchal and later Isaye were fond of quoting St. Thomas’s observation: “It is self-evident that truth exists, for he who denies its existence, grants that truth does not exist; for, if truth does not exist, it is true that truth does not exist; but if something is true, then truth must exist” (Maréchal, 1970, p. 90; quotation from St. Thomas, *Summa Theologica*, I, 2, 1, ad 3). St. Thomas called this type of clearly self-referential argument ‘*redargutio elenchica*’. This “method of discussion” (*procédé de discussion*) Isaye called ‘*rétorsion*’ (Isaye, 1954, p. 205).

‘*Rétorsion*’ in French has the general meaning of responding to someone using the same method that he or she used against oneself—or, in other words, retaliation, but the method came to have a more specific meaning as Isaye defined it. In what follows, I therefore will keep to the French spelling of ‘*rétorsion*’ as a reminder that this term refers to Isaye’s particular method of argumentation.<sup>70</sup>

*Rétorsion* proceeds in the following way:

[C]ertain objections are made in such a way that he who objects, by the very fact of his objection, by the act of its exercise, concedes the thesis that he wished to deny or place in doubt. By directing the attention of the objector to the concession that he has implicitly made is to swing the objection to my favor, this is to retort, to make a *rétorsion* [*faire une rétorsion*]. (Isaye, 1954, p. 205)<sup>71</sup>

Thomas’s *redargutio elenchica* was an informal method that consisted in demonstrating that the content of an opponent’s assertion is inconsistent with the very *fact* of his asserting what he does. —So far, this of course looks very much like pragmatistical self-referential argumentation that we’ve seen before. However, Isaye sought to take the *redargutio elenchica* of Thomas and de-

---

<sup>70</sup> Donceel (1974) has followed this practice, using the word ‘retorsion’ with an ‘s’, similar to the Latin *retorsio*. One of the rare published papers in English about Isaye, by Martin X. Moleski, preferred to use ‘retortion’ with a ‘t’, as he explained: “I prefer the alternative spelling given by the *Oxford English Dictionary* because this calls attention to the cognate, ‘retort’. The *Q.E.D.* indicates that ‘retortion’ was in use as early as 1610 to refer to ‘an answer made to an argument by converting it against the person using it’ ” (Moleski, 1977, p. 61). Because of the specific meaning Isaye came to associate with the name of his method, my preference is to stay with his French term.

<sup>71</sup> This and subsequent translations of Isaye’s work are mine.

velop it into a method of philosophical argument with a capacity to demonstrate “first truths [that have] an objective character (not relative to the subject), [that are] necessary and unchangeable” (p. 219). This motivation, in itself, places the goal of Isaye’s *rétorsion* beyond that of pragmatical self-referential argument. It is of course a different question whether Isaye was successful.

He claimed that *rétorsion* is *not* an *ad hominem* argument:

*Rétorsion* is not a simple argument *ad hominem*.... It is not a matter of turning back a particular thesis (true or false) against an adversary, a thesis that he otherwise admits: Another opponent to our first truths would not be touched by such an argument. *Rétorsion* takes the adversary *as adversary*. He expresses an objection, and that suffices to give rise to the *rétorsion*. (p. 218)

What Isaye is pointing to are the *inevitable* commitments, in his view, that *anyone* has already made who engages in rational discourse—anyone, not just the person who *in fact* happens to be involved in a philosophical exchange. “The necessity of [such a] commitment is a guarantee of objectivity. What is binding is an objective truth” (p. 219). Although for Isaye *rétorsion* is implemented in a factual context, in the context of dialogue with another, the method is intended to identify “first truths” that are necessary and invariant (here I substitute my own term which I think expresses Isaye’s meaning) no matter who makes an affirmation.

“*Rétorsion* is exercised against the hypercritical [i.e., skeptical] adversary who would deny one first truth or another” (p. 209). These “first truths” appear in Isaye to resemble Passmore’s invariant conditions of discourse. Isaye gives an example from Aristotle: “*Impossibile est eidem simul inesse et non inesse idem secundum idem*” (p. 206)—that it is impossible for a thing to both be and not be in the same respect. Later, Isaye gives the examples of the principle of non-contradiction and the principle of objectivity. Any attempt, *by anyone*, to deny an alleged first truth of this kind results, according to Isaye, in a self-referential inconsistency that is both necessary and invariant with respect to *any* asserted similar denial.

Again, it is not my purpose here to evaluate Isaye’s method, but rather to identify it as an approach which he intended not as a form of *ad hominem* argument, nor as a kind of pragmatically self-referential argument. As he saw it, the purpose of his method of *rétorsion* is to bring to light those absolutely binding commitments, without which reason and knowledge are, in principle,

not possible—here lies the Kantian-transcendental aspect of his thought. As Isaye commented on the application of his method of *rétorsion*: “[A] simple transposition of vocabulary, using a small dictionary easy to establish, would transport our complete proof, without affecting its strength, to Kant’s transcendental horizon....” (Isaye, 1953, p. 73).

Let us return to the hypercritical adversary. For him to pose his objection he has had implicitly to concede first truths. But he has not noticed this. Why? ... The skeptic is hypnotized by what he affirms in words, by what he specifically intends [*actu signato*]; he forgets the *act exercised* [*l’acte exercé*].... (Isaye, 1954, p. 220) And how to lead the adversary to reflect, to bring him to affirm in his own words what he has already affirmed on an experiential level...? Here enters the art of *rétorsion*. It is not a matter of employing a mechanical procedure. Each time it is necessary to adapt to a new situation, to a new vocabulary.... (p. 221) The philosopher who retorts needs always to choose a manner of expression that is faithful to the thought of the objector. He needs to ask the objector if his objection is objectively expressed. *Rétorsion* is based precisely on the admission of the objector: “That is my thought. This is what is true.” (p. 222)

Notice in this passage that Isaye refers to *rétorsion* as “an art,” one that needs to be sensitive and responsive to the manner in which a philosophical opponent thinks. To accomplish this in a way that provides objectivity, he suggests writing down what the opponent affirms and then discussing the objectively written formulation with him in order to achieve a kind of meta-level exchange (not Isaye’s terminology) with the other person (Isaye, 1954, p. 223). Then it may be possible to analyze the written formulation in a more detached/objectified manner.

*Rétorsion*, in short, is an individually applied art/technique that seeks to encourage a philosophical opponent to reflect upon his position, to examine it in a detached way, to recognize that in his very act of affirmation he necessarily is bound to underlying commitments, to first principles, essential to the very capacity to exercise reason and to achieve objective knowledge. “Any concrete judgment whatever, no matter how contingent it may be, poses first principles (not in its verbal expression, but through the exercise of the activity of affirmation)” (Isaye, 1953, p. 45).

To accomplish this series of goals, *rétorsion* functions, then, as a rhetori-



cal method, which, as we've seen earlier in this study, places the primary focus upon *persuasion*, upon interpersonal exchanges that will ideally lead to an opponent's change of view. And yet, as Johnstone and also Passmore have made clear, rhetorical argumentation—even when it is undertaken by means of a method that cannot *not* be accepted without undercutting the very possibility of rationality—is always a shaky affair, encumbered by the “almost inconceivable hardihood” (Passmore) of an opponent's allegiance to his own beliefs, coupled with what in the opponent may be “an almost unlimited capacity for sophistry” (Johnstone).

Where philosophical commentaries relating to Passmore's work have sometimes misrepresented his approach to absolute self-refutation as pragmatically self-referential in nature, some commentaries about Isaye have mistakenly identified his method of *rétorsion* with *ad hominem* argument, despite Isaye's explicit repudiation of this label, as quoted earlier. For example, Moleski (1977, pp. 61-62) wrote:

Retortion is essentially a process of recognizing inconsistency in a philosophical position. It results in the judgment that no person could adopt such a position without becoming involved in a kind of self-contradiction. This places it in the genre of *ad hominem* arguments, although “the *Homo* in question is every *Homo*, every human being.” [Quoting Donceel 1974, p. 81].

Moleski continued: “*An argument which is subject to retortion is rejected because no one can adopt it consistently, not simply because the argument is inconsistent with a particular person's beliefs.* Since it is implicitly concerned with all men, retortion can lead us to *a universally valid statement...*” (Moleski, 1977, p. 62, italics added). The latter correct description of Isaye's *rétorsion* places it beyond what philosophers understand as *ad hominem* argument. A more accurate description was given by Muck (1968, p. 173):

[T]ranscendental retorsion is ... not simply an *argumentum ad hominem*, which refutes a statement by showing that it contradicts another statement accepted by the opponent. Transcendental retorsion tries to show that the very statement against which the objection is raised must, at least implicitly, be recognized in order to be able to raise the objection.

It may also be tempting to regard Isaye's *rétorsion* as a pragmatically self-referential method, given that the method is applied both to the factual expression by an opponent and to the fact of the opponent's implicitly affirmed, rationally binding commitments. But this, also, would be a mistake. As Jean Ladrière pointed out:

One might think of allying the method of *rétorsion* with analyses of the pragmatic conditions of language. Pragmatics distinguishes, for any utterance, the content of the utterance (expressed by a proposition) and the act of language that is brought about by the fact that this content is effectively expressed. That act has specific conditions of realization. A statement that denies one of its own conditions of enunciation would be a pragmatical contradiction. *Rétorsion* consists, in a certain sense, in disclosing a pragmatical contradiction. If the contradiction is untenable, that which was denied appears as incapable of being denied, therefore as necessarily valid. This association of the method of *rétorsion* with that of pragmatics cannot, however, authorize a pragmatical reinterpretation of *rétorsion*. Because the point of view of pragmatics remains that of an exterior analysis. It is not the act itself that is disclosed, but another act which, from the outside, takes apart the mechanism involved. [In contrast] it requires the disclosure of the *reflexive* dimension of the act in order for an argument by *rétorsion* to gain its full effectiveness. (Ladrière, Préface in Isaye, 1987, p. 20)<sup>72</sup>

Ladrière's point here seems to be that the frame of reference required in pragmatical argumentation is *external* to the statement or utterance being considered; *rétorsion*, in contrast, is essentially an *internal* affair: It relies upon the reflexive capacity of both individuals who are involved in dialogue to recognize that a self-referential inconsistency has occurred, but this recognition does not take place from outside the opponent's framework (as when it is explicitly pointed out to an adversary that he has, in fact, done the very thing he denies); instead, *rétorsion* relies upon the adversary's capacity to see that he has implicitly breached his own fundamentally binding commitments. In short, *rétorsion* is an essentially self-referential method to which the opponent succumbs because it is actively exercised internally and reflexively by him himself.

---

<sup>72</sup> My translation.

...

I have devoted a fairly long discussion to Isaye's method because it is a good deal less familiar to Anglo-American philosophers than Passmore's and Mackie's approaches. As was the case with Passmore and Mackie, Isaye's method leaves one with important unanswered questions. His concept of "first truths" is in need of clear definition: Are the "first truths" that *rétorsion* is capable of bringing to light actually the Kantian transcendental preconditions required in order for objective knowledge to be possible? Isaye gave only a small number of examples of such truths. There is a need to spell out in detail how we are to know, and how we are to justify in a rigorous fashion, that such truths are, as he claimed, "necessary and unchangeable."

Nonetheless, it is evident that by using self-referential approaches to argumentation, both Isaye and Passmore sought to develop methods whose similar purpose is to identify principles or truths that function as invariant preconditions of discourse or rationality. Both philosophers' approaches are "*activity-based*," focusing upon the performance-based conditions that underlie factual assertions or affirmations. There is therefore an evident rhetorical—that is, persuasive—goal of their respective methods: Both Passmore's approach to absolute self-referential argumentation and Isaye's method of *rétorsion* are intended to be applied in an interpersonal context of philosophical dialogue where the objective is to be able to argue with an adversary in a way that will ideally produce rational conviction in him or her. However, as Passmore and Johnstone have recognized, this is not always easy, and in fact it can be a nearly impossible challenge when confronted by opponents who are deeply committed to their own allegiances and will attempt in whatever ways they can to dodge damaging criticism. In this, Isaye's and Passmore's self-referential approaches sink or swim as a function both of the skill of the self-referential analyst and of the human variability of those to whom their arguments are addressed. From this point of view alone, rhetorical argumentation and the conceptual variety are inherently distinct, both in their ends as well as in the degree of success to which they can attain in reaching them.

In studying the work of Passmore and Isaye, and to an extent also that of Mackie, I have come to understand what may be their *intended* although not fully realized contributions to philosophical self-referential argumentation: Both Passmore and Isaye, as I read their work, sought to develop approaches to self-referential argumentation that are neither *ad hominem* nor performatively self-referential, but rather, in a broader, theoretically more general sense, comprise transitional, not fully formulated approaches situated between

strictly pragmatical argumentation and the metalogical variety, to which I now turn.

### 6.8 Metalogical self-referential arguments

[W]e are brought to the conclusion that we can never transcend the limits of possible experience.

– Immanuel Kant (1965/1929, B xix)

It constitutes a great advance in our critical attitude ... to realize that a great many of the questions that we uncritically ask are without meaning.... [O]ne is making a significant statement about his subject in stating that a certain question is meaningless.

– P. W. Bridgman (1961/1927, pp. 28-29)

Unlike strategies of argumentation using pragmatical self-reference, a metalogical approach directs attention to the conceptual commitments that are necessarily involved if a concept *in principle* is to permit reference to those objects to which reference is presupposed. Whereas philosophical argument that relies upon principles of pragmatical self-reference is *rhetorical* and in many cases appropriately considered to be *ad hominem* in nature, self-referential argumentation developed on a metalogical basis has an unmistakable *conceptual* and *transcendental* focus.

Metalogical self-reference is distinct from pragmatical self-reference in another, theoretically fundamental respect that we shall explore later in some detail: Pragmatically self-referentially inconsistent statements generally undermine their own truth, and therefore are regarded as self-falsifying.<sup>73</sup> In such a case, there is, so to speak, a *truth-functional short-circuit*. Metalogically self-referential inconsistencies, in contrast, undermine their own capacity *in principle to be meaningful*. As Quine was quoted earlier, “reference is *nonsense* except relative to a coordinate system” (Quine, 1969, p. 48, italics added). If, for example, I claim, “I can refer to what is beyond my capacity to refer,” I have made an assertion that is without possible sense.

There is a necessary interconnected relationship of logical dependence between object identification and the system of reference that makes this pos-

---

<sup>73</sup> “[T]hey are false because they are inconsistent with the facts that are given in and by any assertion of them. Thus they are not in themselves self-refuting, but to try to assert any of them is self-refuting” (Finnis, 2004, p. 13).

sible. As noted in {5.7–5.8}, the identity of a set of objects of reference is inseparable from, and logically *can have* no meaning apart from, the reference system that makes reference to them possible. Where pragmatical self-referential arguments attempt to show that a speaker's statement falsifies itself, metalogically reflexive arguments that are negatively critical seek to demonstrate that a given concept undermines its own capacity in principle to refer, and in the conceptual step to accomplish the impossible, it undermines its capacity to be meaningful.

In what follows, it will often be convenient to use language that places the particular concepts under analysis within the context of a theory, a position, or claims in which those concepts play a role. This is a matter of expressive convenience: As it should be clear by now to the reader, our interest will be in the referential preconditions required by the *concepts* studied, not in the factual conditions of particular *assertions*. The contingent uses to which such concepts may be put do not constitute the subject-matter we have in view; they are no more than the specific ways in which such concepts are applied in a particular context.

Universally, for a theory, a position, or a claim to function as such it must be capable of referring to certain objects, about which assertion is possible. Metalogical reflexivity comes to be of interest in connection either, from a negatively critical point of view, with theories or claims that conflict with their own referential preconditions, or, from a constructive point of view, with theories or claims that compel assent, since they cannot be denied without producing such a conflict.

A number of past applications of self-referential argumentation lend themselves to classification as metalogical applications of self-reference. I list some of them here so that the reader may gain an idea of the breadth of philosophical issues to which metalogical self-referential argumentation may be applied; we'll have occasion to examine a number of these topics in greater detail later in this study. Such arguments have attempted to identify a wide range of self-undermining concepts and claims. Among these are:

- ◆ Descartes' methodologically skeptical hypothesis (presented as potentially true in reality) of an evil genius, an hypothesis capable of shaking all confidence in our abilities to ascertain the truth about reality;<sup>74</sup>

---

<sup>74</sup> Bouwsma (1965) and Bartlett (1988, pp. 221-232).

- ◆ Kant's distinction between objects spatially structured by the human mind and "objects themselves," to which the human concept of space does not apply;<sup>75</sup>
- ◆ The hidden variable interpretation of quantum mechanics, which expressed a bias in favor of realism and physical determinism on the level of small-particle interactions;<sup>76</sup>
- ◆ Philosophical skepticism as treated by P. F. Strawson;<sup>77</sup>
- ◆ The argument (which ironically depended on a pragmatically reflexive strategy) attempting to show that the rejection of free choice is self-falsifying, or else pointless;<sup>78</sup>
- ◆ The view claiming that solutions to mathematical or other problems are "discovered"; they are not "invented";
- ◆ The opposing view, claiming that solutions to mathematical or other problems are "invented"; they are not "discovered";<sup>79</sup>
- ◆ The doctrine that there exists (or does not exist) a "meta-physical self";<sup>80</sup>
- ◆ The belief that a phenomenological description of an experience tells us what was "already present" in the experience pre-reflectively and implicitly;<sup>81</sup>
- ◆ The Newtonian concepts of absolute time and space;<sup>82</sup>
- ◆ The realist view that accords a separable existence to past or future events, independently of the present;<sup>83</sup>

---

<sup>75</sup> See Bartlett (1988), also Bartlett (1970, Chapter 2.1).

<sup>76</sup> Cf. Bartlett (1980, section VII).

<sup>77</sup> Strawson's argument against skepticism in Strawson (1959) can, with some modest stretching, be interpreted as an attempt to show that the skeptic's position is metalogically self-undermining.

<sup>78</sup> This argument was advanced in Boyle, Grisez, & Tollefsen (1976). Although a hard-working attempt to show that freedom of choice may be rejected only on pain of pragmatic self-referential inconsistency or pointlessness, the argument itself is metalogically self-undermining. See Bartlett (1979).

<sup>79</sup> On this hypothesis and the preceding one, see Bartlett (1978a, pp. 70-72, 79-82).

<sup>80</sup> Bartlett (1970, Chapter 2.6), also Bartlett (1978b).

<sup>81</sup> Bartlett (1975a, section III) and Bartlett (1974).

<sup>82</sup> Bartlett (1970, Chapter 2.1).

<sup>83</sup> Bartlett (1970, Chapter 2.1).

- ◆ The framework-independent concept of absolute truth,<sup>84</sup>
- ◆ The doctrine claiming that every event is the effect of a prior cause, and the related doctrine claiming that in a cause-effect sequence, the occurrence of the cause was indispensable to the occurrence of the effect;<sup>85</sup>
- ◆ The interrelated beliefs that there is a common “pole,” called ‘the ego’, shared by all of the investigator’s experiences; that consciousness is a universal attribute of experience; that consciousness is a kind of “container” of experiences, beyond which meaningful claims may be made,<sup>86</sup>
- ◆ the doctrine that mental events are in many instances the results of prior acts (a belief inspired by the causal dogma mentioned earlier);<sup>87</sup>
- ◆ The belief that reflection does (or does not) perturb the structure or nature of pre-reflective experience,<sup>88</sup>

### 6.9 The constructive use of metalogical self-reference

The constructive use of metalogical self-reference depends upon a special property of claims of a certain kind: This is the property possessed by a claim that is such that its denial leads exactly to the variety of self-referential inconsistency that is in view here, i.e., self-referential inconsistency that *precludes* that the intended reference of the claim is possible at all.

As mentioned earlier in passing, claims of this kind are called ‘*self-validating*’, since, if they are rejected, they succumb to self-referential inconsistency of such magnitude that *their capacity to be meaningful is undermined*, a phenomenon we shall subsequently look at more closely.

As in the case of pragmatically reflexive arguments, there is an interplay between the critical and the constructive ends to which metalogically reflexive arguments may be put. The relation between criticism and construction is similarly bridged here by a conditional: If it can be shown that a claim is metalogically self-undermining, then the rejection of that claim will compel assent. It is important to notice that the *rejection* of such a claim does not en-

---

<sup>84</sup> Bartlett (1970, Chapter 2.4).

<sup>85</sup> Bartlett (1970, Chapter 2.5).

<sup>86</sup> Bartlett (1970, Chapter 2.6).

<sup>87</sup> Bartlett (1970, Chapter 2.7).

<sup>88</sup> Bartlett (1970, Chapter 2.7).

tail the positive endorsement of its negation. For example, the rejection of “there exists a metaphysical self” does not commit us to “a metaphysical self does not exist.” —Both claims employ a *framework-transgressing concept* that stands in conflict with its framework-relative basis. We shall also explore this in detail later.

Among positions and arguments that have sought their own validations in ways closely akin to a metalogically reflexive strategy, again to give the reader some representative samples, these could be listed:

- ◆ Kant’s transcendental deduction;
- ◆ Collingwood’s absolute presuppositions of systematic thought, which in his view are presupposed by any cognition, and make knowledge possible;
- ◆ Husserl’s conception of transcendental phenomenology, the analysis of which reflexively discloses the necessary foundation for its own possibility;
- ◆ Strawson’s attempt to deduce, in a quasi-transcendental manner, the necessary and basic structure of a conceptual system that makes objective knowledge possible;
- ◆ Gaston Isaye’s transcendental method of *rétorsion*, which seeks to identify the conditions of the possibility of reason and of objective knowledge;
- ◆ The following pair of mutually reinforcing positions: The author’s reflexive argument that metalogical referential consistency is a necessary condition of meaning, on the one hand, and his relativistic theory of the constitution of experience, on the other. Together, these approaches show that a wide range of everyday and technical concepts is metalogically self-undermining, underscoring the need for a vocabulary of radically different but referentially self-consistent concepts.<sup>89</sup>

---

<sup>89</sup> Cf., *inter alia*, Bartlett (1982) and Bartlett (1970).



# 7

## Possibility Theory

The domain of the possible plays a prominent part in our thought about the affairs of nature and of man. Deliberation about alternatives, contingency planning, reasoning from hypotheses and assumptions, and thought-experiments are but a few instances of our far-flung concern with possibility. The rational guidance of human affairs involves a constant recourse to possibilities; we try to guard against them, to prevent them, to bring them to realization, etc. The theory of possibility thus represents a significant part of our understanding of man's ways of thought and action.

– Nicholas Rescher (1975, p. 1)

### 7.1 The basic vocabulary of the metalogic of reference

Certain of the principal purposes of the metalogic of reference are shared with transcendental epistemology as it has evolved since Kant: In that tradition, the ideal objective of transcendental argumentation has been to identify, and then to justify, invariant preconditions that underlie the possibility of knowledge. This goal is evidently highly general and abstract, and it has been pursued by philosophers in a multitude of often divergent ways, many of whom have been Continental philosophers, and some Anglo-American. The multiplicity of proposed approaches, ranging from phenomenological to analytic, resists any characterization of transcendental epistemology in a clear and well-defined way. Too, none of these approaches to transcendental epistemology has been developed by means of an explicit study of the conditions and principles that are the foundation of *all reference*, and derivatively, of all so-called 'referring activity'. For these reasons, I have chosen not to subsume the metalogic of reference under traditional approaches to transcendental argumentation, despite sharing in some of its goals.

The bare and informally stated schema of the transcendental objective is, as I've mentioned, to identify, and then to justify, invariant preconditions that

underlie the possibility of knowledge. To accomplish this in any systematic and logically persuasive way requires the development of a clear understanding of a specific set of conceptually fundamental concepts, foremost among them, the concept of possibility and the related concept of necessity, combined with the concept of transcendental preconditions, which are usually thought of as a special variety of presupposition. A well-developed theory of possibility has yet to be both formulated and widely accepted; the same is true both of the concept of presupposition and of the particular variety of presupposition that has been given the role of transcendental precondition. Beyond the task of meeting these particular needs, a rigorously formulated approach to transcendental justification remains a distant project.

In this chapter, I consider the modal concepts of possibility and necessity; in the next chapter, the concept of presupposition; and then in the next chapter, I examine the logic of transcendental argumentation.

## 7.2 The concept of possibility

The history of the concept of possibility would require a book in itself. Instead, for our purposes here we'll look ahistorically at a group of concepts of possibility that can be placed along an approximate spectrum in terms of which they can be ordered, from minimally abstract and with a lesser scope of application, to those concepts of possibility which provide a more comprehensive understanding of the fundamental meaning of the idea of possibility, and which gradually approach the level of maximum theoretical generality that the metalogic of reference requires. The perspective we acquire through this review of concepts of possibility will help to situate in relation to the range of concepts of possibility the theory of possibility that I subsequently formulate, one that meets the specific requirements of a metalogical approach to transcendental argumentation.

## 7.3 The spectrum of possibility

### 1. *Psychologically based possibility*

This is one of the most basic and intuitive concepts of possibility. There are several ways in which the concept of possibility can be psychologically based. The most familiar to philosophers is possibility understood as psychological conceivability. There are two senses in which psychological conceivability may be understood: Possibility can be equated with what a particular human mind is capable of conceiving, which in essence limits possibility to what is

imaginable by that individual mind; or the notion of possibility can be broadened in an attempt to apply to all human minds. In either case, human psychological limitations will delimit the resulting range of possibilities that can be recognized.

The range of possibilities that are psychologically conceivable has often been thought to be restricted to what is not self-contradictory, but some researchers in contradiction-tolerant and in contradiction-affirming logics (also called ‘paraconsistent’ and ‘dialetheic’ logics) appear to embrace the view that, at least for some of these philosophers, contradictions are psychologically conceivable.

Psychological abilities—for example, the ability to speak French, or the ability to remember how to factor and use trig functions—have also been recognized as comprising a variety of psychologically based possibility. “It is possible for him to speak French, although he is not speaking French now” and similar statements express this concept of possibility.

## *2. Temporally based possibility*

Here, too, alternative concepts of possibility have been proposed. Diodorus Cronus of Megara suggested that possibility should be understood as what is true or will be true, so that the range of the possible is coextensive with truths in the present or the future. Alternatively, among the Stoics, Chrysippus proposed several views of possibility, among them that possibilities should be defined as that which is sometimes true, and necessity by that which is always true.

In modern times, logician Jan Łukasiewicz suggested a novel temporally based concept of possibility according to which the range of the possible is determined by past forgotten events and their exhausted consequences:

If, of the future, only that part is real today which is causally determined by the present time ... then also, of the past, only that part is real today which is still active today in its effects. Facts whose effects are wholly exhausted, so that even an omniscient mind could not infer them from facts happening today, belong to the realm of possibility. We cannot say of them that they were but only that they were possible. And this is as well. In the life of each of us there occur grievous times of suffering and even more grievous times of guilt. We should be glad to wipe out these times not only from our memories but from reality. Now we are at liberty to believe that when all the consequences of those fatal times are

exhausted, even if this happened only after our death, then they too will be erased from the world of reality and pass over to the domain of possibility.<sup>90</sup>

### 3. *Physical or nomological possibility*

In this view, the domain of possibilities is constrained by physical laws: What is possible is what physical laws permit. If currently accepted physical laws do not permit a certain occurrence, process, or event, then it is considered, at present, to be physically impossible. “A human being can travel from Earth to Alpha Centauri in 30 minutes of elapsed Earth time” is at present physically impossible. In this sense, possibility is constrained by the recognized physical structure of nature.

### 4. *Epistemic possibility*

Related to the preceding concept of possibility is the view that, relative to our present knowledge, certain claims are possible (possibly true), and others not. In this view, a certain state of affairs is judged to be epistemically possible if it is not excluded or prohibited by what we now know. Relative to current knowledge, if to assert such a state of affairs is inconsistent with what we now know, then that state of affairs is judged to be epistemically impossible. If we do not know that something is not the case, then it is epistemically possible.

### 5. *Formal possibility*

Recurring in the views of many philosophers and logicians, possibility has been associated with the absence of formal, logical contradiction. Possibility understood as freedom from self-contradiction has a long heritage, proposed in a variety of ways from Aristotle to Aquinas, Descartes, Spinoza, Leibniz, and Kant, to modern times. The proliferation of systems of formal logic has made it clear that formal possibility is systems-relative: freedom from contradiction can be provided by formalized systems in numerous ways; relative to each can be associated a range of formal possibilities.

There are variations on this view: From the standpoint of a *consistency theory of possibility*, the consistency of propositions in relation to one another, or the self-consistency of a proposition with itself, may be equated with what is possible. The limits of relations of consistency among propositions, relative to one or a set of formal systems, determine in this approach the boundaries of the possible.<sup>91</sup>

---

<sup>90</sup> Quoted in Prior (1967, p. 28).

<sup>91</sup> Fitch (1950b), e.g., defined possibility in terms of “self-consistency”—“A proposition is

### 6. *Possibility as complementarity*

During the past several decades, inconsistency-tolerant and inconsistency-asserting approaches to mathematical logic have been developed. Such systems loosen the classical constraints upon formal logical possibility, recognizing the potential legitimacy of such statements as ‘ $A \wedge \sim A$ ’ as admissible expressions, subject to qualifications and limitations depending upon the system. The concept of possibility expressed by systems permitting or asserting contradictions forms a variety of possibility that is clearly subordinate within the category of formal possibility. But it may be of greater explanatory value to recognize such a concept of possibility as comprising a distinguishable form of possibility, given the distinctive capacity of such systems to allow for conflicting propositions both of which may be asserted. Such systems may be considered to express “complementarity” in this sense.

In a parallel fashion, in the Copenhagen interpretation of quantum mechanics (see {27}), quantum particles may, e.g., be claimed to exist not in one state or another, but in a sense in all possible states at once, and in this context complementary, but classically conflicting, statements can be predicated of them. This concept of possibility could appropriately be subsumed under physical/nomological possibility, but to recognize its distinctive character through the allowance of complementary assertions, it is here placed in its own category of possibility.

In more general terms, a physical object or event, or a formally represented proposition, about which logically incompatible statements may be applied, presupposes a concept of possibility as complementarity.<sup>92</sup> As we

‘possible’ if and only if it is logically self-consistent” (p. 370)—but without explicitly defining the term ‘self-consistency’. In Fitch (1952), he instead defined ‘consistency’ in terms of possibility: “The concept of *consistency*, as applying to propositions, can be defined in such a way that to say that  $p$  is consistent with  $q$  is to say that the proposition  $p \ \& \ q$  is (logically) possible” (§12.23, p. 75).

<sup>92</sup> A related view was first implicitly proposed by the Russian, Nikolai A. Vasil’ev (1880–1940), who suggested the general idea of a three-valued logic in four papers published between 1910 and 1913. Kline (1965, p. 320) summarizes Vasil’ev’s idea as foreshadowing a complementarity concept of possibility: “In Aristotelian logic, ... a given object cannot provide the ground for both affirmative and negative propositions, since ... ‘ordinary’ Aristotelian negation is based on incompatibility. But a non-Aristotelian logic, without the Law of Noncontradiction, would not utilize such ‘ordinary’ negation.” According to Vasil’ev, we can imagine a world in which “... empirical fact  $a$  would make the proposition ‘ $S$  is  $A$ ’ true; fact  $b$  would make the proposition ‘ $S$  is not  $A$ ’ true. But facts  $a$  and  $b$  would not be incompatible.... [A] given sense object might thus simultaneously ground both affirmative and negative propositions about the sensed properties of the object.”

On the possibility of such “realized contradictions,” cf. Comey (1965, p. 368) and Vasil’ev (1924, p. 108). See also Wittgenstein’s interesting speculations in his (1956, Part V), also

shall see later on, a modal law of non-contradiction may need to be presupposed:

$$\sim(\diamond p \wedge \sim\diamond p).$$

What is free of contradiction in the latter sense remains possible within the complementarity view.

### 7. *Parametric possibility*

This conception of possibility has not attracted much attention despite its potentially wide range of applicability. It was advocated by philosopher Scott Buchanan (1927) in his book *Possibility*, a published version of his doctoral dissertation and seldom mentioned in later literature. Buchanan described possibility in parametric terms: As in mathematics, a parameter is an identity condition that specifies a field of variability within which values are related according to a rule:

[In parametric formulae] one can discriminate (1) what I shall call an identity condition or constant, (2) a class of particulars ... called the field of variability, and (3) a rule of order, or set of relations ... which holds between the particular determinations or members of the class. These are the three phases of any parameter, or we may say a parameter has an identity condition, a field of variability, and a rule of order. (Buchanan, 1927, pp. 37-38)

In  $f(x)$ ,  $x$  varies according to conditions fixed by  $f$ . In the equation  $x + 5y = k$ ,  $k$  determines a family of lines. When  $k$  assumes a specific numerical value, the equation identifies a specific straight line. In the equation,  $x$ ,  $y$ , and  $k$  are of course all variables, but Buchanan recognized their role as parameters:  $k$  is a parameter of higher order that delimits the family of lines which, in the given equation, the values of lower order parameters  $x$  and  $y$  serve to identify.

In Buchanan's view, the parametric concept of possibility is therefore hierarchical: In the equation above,  $k$  functions as a blank field of variability that is delimited by the rules set by the equation involving  $x$  and  $y$ , and in which  $k$  has a higher-order governing role.

The concept of possibility that emerges from this kind of analysis is at once highly abstract and general. A parametric understanding of possibility

sees possibility in terms of rule-defined and rule-delimited variability. Buchanan's generalized notion of parameter is a way of articulating what such rule-based variability means. Possibilities are therefore parametrically delimited; when the limits of variability of parameters are violated, we encounter impossibilities. But how this happens and how we are to recognize such violations in individual instances, and therefore gain a clear conception of parametric possibility, Buchanan did not spell out.

For many readers, Buchanan's parametric concept as it stands does not provide an intuitively transparent or easily employed concept of possibility. Buchanan did not develop his abstract concept of possibility built upon the mathematical idea of parameter to a stage where it can be directly and effectively applied by philosophers and logicians. And yet it is, because of its generality, a potentially integrative and broad concept of possibility.<sup>93</sup>

### 8. *Many worlds possibility*

Of the concepts of possibility identified in this chapter, the many worlds concept of possibility has at the present time become the most popular and widely discussed among philosophers and logicians. The general notion of "possible worlds" is associated historically with Leibniz, who suggested that there are as many possible worlds as things that can be conceived without contradiction, but he apparently did not make the further step, now commonly made, to claim that necessary truths are true in "all possible worlds."

The many worlds conception of possibility began to invite serious attention during the 1960s. The modal concepts of necessity and possibility were semantically interpreted as quantifiers over possible worlds, according to the two general principles: Using the symbols ' $\square$ ' and ' $\diamond$ ' to stand for 'necessity' and 'possibility', respectively:  $\square\phi$  iff (if and only if)  $\phi$  is true in every possible world,  $\diamond\phi$  iff  $\phi$  is true in some possible world, and  $\sim\diamond\phi$  iff  $\phi$  is false in all possible worlds.

The notion of possible worlds naturally leads to several questions: How is a "possible world" to be defined?, How is one possible world to be differentiated from another?, and What does it mean for something to exist/be included in a possible world?

Numerous answers and interpretations, some widely divergent, some conflicting with one another, and varying greatly in their informative value, have

---

<sup>93</sup> My discussion of Buchanan's view is necessarily condensed here. It bears mentioning that later in his book he sought to extend his parametric view in terms of whole/part relations. For example: "Possibility is the regulative idea for the analysis of wholes into parts. Parts are then possibilities or potentialities with respect to their respective wholes and systems are hierarchies of such possibilities" (Buchanan, 1927, p. 81).

been offered to these questions; they are answers and interpretations that will not concern us here. Collectively they may be thought to comprise ways that have been proposed in terms of which reference is possible, or is claimed to be possible, to the set(s) of objects of reference that is (are) said to populate or make up, and differentiate, such possible worlds.

### 9. *Framework-relative possibility*

Apart from the imagination-catching, science fiction cachet of the phrase ‘possible worlds’, in my own perhaps against-the-grain estimation, little of theoretical value is gained by talk of such “worlds” if our interest is specifically in understanding the nature of possibility. Instead, in a sober and maybe less exciting conception of possibility, a possible world may broadly and informatively be understood as a framework of reference permitting a range of object identifications, where the nature of such “objects” stands then in need of specification. For readers whose interest lies in the notion of possible worlds, the fundamental general question they may wish to answer is, what criteria must such a frame of reference satisfy in order for it to substitute for or fulfill the intended role of a “possible world”?<sup>94</sup>

A liberal and inclusive answer, which I shall accept here, is that any frame of reference that provides a system in terms of which object identification can take place may be considered a “*possible world*.”<sup>95</sup> —But, once again, as I see this, to characterize such a frame of reference as a “possible world” is not informative, for defining possibility in terms of possible worlds begs the question by shifting to a concept in which possibility is already assumed to be embedded: When it is believed that the concept of possible worlds sheds light on the concept of possibility we immediately notice the circularity involved, a circularity common in much possible world theorizing.<sup>96</sup> If our interest is in developing an understanding of what, at a genuinely fundamental theoretical level, constitutes possibility, we are not well-served by passing over this issue. I therefore do not consider possible worlds to offer

---

<sup>94</sup> A wide variety of answers has been proposed, including the notion that possible worlds are distinguished by virtue of: not being spatio-temporally overlapping (e.g., Lewis, 1986); or being consistent states-of-affairs (e.g., Plantinga (1974, 1976); or being physically consistent recombinations on the atomic level of the actual world (e.g., Armstrong, 1986); etc.

<sup>95</sup> Note that there is no stipulation made in this open-ended answer as to the sort of objects that may be included, and no stipulation is made concerning the degree of specificity or uniqueness of such identification.

<sup>96</sup> To give an example: A proponent of possible worlds, David Lewis, proposed, without apparent concern over such circularity, that a *possible world* is an actual maximally consistent representation of how the universe could *possibly* have been (Lewis, 1986). Defining ‘possibility’ in terms of what is possible hardly informs us what ‘possibility’ means.



significant illumination given the interest here in explaining what, at its most fundamental level, possibility means.

Despite this reservation, the semantic metaphor of possible worlds has thrown some light on the meaning of the traditional broad notions of “absolute” and “relative” possibility<sup>97</sup> by distinguishing, on the one hand, the absolute case, in which a proposition may be true in all possible worlds (or as in the present study: in all reference frames), where every world is possible in relation to every other, and, on the other hand, the relative case, in which, from the standpoint of some possible worlds (or reference frames), only certain others are “accessible,” that is, possible. This approach has led to much discussion about “accessibility relations” between possible worlds, relations which, in other words, have to do with whether a proposition may be true in all possible worlds (that is, true no matter what frame of reference is employed), or whether that proposition may be true in a more limited, relative way, that is, true in every world that is accessible from a given world (true only in reference frames to which a given frame provides the basis of reference).<sup>98</sup> In the many worlds view, to be possible means to be true in a possible world—i.e., true in a framework of reference.

As we shall see in subsequent chapters, the three previously mentioned central questions often raised about the concept of possible worlds—How is a possible world to be defined?, How is one possible world to be differentiated from another?, and What does it mean for something to exist/be included in a possible world?—are more easily, directly, and informatively answered when we deal, instead, with the concept of frameworks of reference.

### 10. *Metalogical possibility: The preconditions of identification*

Earlier in this study (see {5}), we saw that the concept of reference is essentially tied to identification, that is, any instance of reference is at once an

---

<sup>97</sup> Traditionally, absolute possibility (also called conceptual or *a priori* possibility) has been equated with whatever is consistent/compatible with basic conceptual necessities (not necessarily formal, logical truths, but consistent with accepted conditions of discourse). In this sense, absolute possibilities include formal, logical possibilities (those that are classically consistent).

Again traditionally, relative possibility has been understood in the sense that a thing may be relatively possible in relation to current technological capacities, or relatively impossible in relation to those capacities. Similarly, a thing may be relatively possible or impossible in relation to known physical laws. etc.

For the traditional distinction between absolute and relative possibility, cf. Lewis & Langford (1932, pp. 67, 445; also 161, 215n, 472, 475f).

<sup>98</sup> For a related, early many-worlds interpretation of relative possibility, cf. Kripke (1963, p. 70).

instance of identification, whether such identification is vague or precise. Any object to which reference is possible *is* an object that possesses an identity. To be a possible object of reference *is*, in other words, to be an entity with some degree of identifiable identity. We noted how frames of reference make it possible to refer to diverse ranges of objects, and thereby to recognize the identities of those objects—to know what is being specified, talked about, described, adumbrated, etc., as distinct from all that is not so specified. I described how frames of reference function in this capacity as systems of coordination in real, non-metaphorical ways, permitting the coordination, for example, of locations with objects; of persons in relation to things; of things and certain times; of persons, things, and times, etc., and recalled Quine's undeveloped remark that "reference is nonsense except relative to a coordinate system." In this sense, the coordinative function of frames of reference lies at the very basis of our ability to know and to communicate knowledge of any class of objects of reference.

This brings us directly to the metalogical concept of possibility: When we ask what is possible, we presuppose that reference to such possibilities as may be in view is somehow assured. The frame of reference that we unavoidably employ, implicitly as background or explicitly in its application, then becomes the object of our reflective attention. The range of possible objects of reference is determined by the capabilities of reference of a given frame of reference: In other words, metalogical possibility is framework-relative possibility, which in turn means identifiable as a function of a particular frame of reference. When we identify the preconditions of reference satisfied by a given frame of reference, we make explicit what is possible from that standpoint.

Suppose we wish to consider a description  $d$  of a certain putative possibility  $p$  to which reference can obtain relative to a certain frame of reference  $f$ . We ask, Is that a bona fide possibility, an impossibility, or a meaningless description? The question, to be answerable, presupposes that the frame of reference in question is associated with a range of possibilities, that is, with a range of possible objects of reference. Should it turn out to be the case that what we assume to be a significant description fails to comply with the referential preconditions of the identification framework in the context of which it is given, then the description fails to refer to a possibility.

In other words, a description  $d$  of a putative possibility  $p$  from the standpoint of a framework  $f$  permitting identifying references is said to identify  $f(d(p))$  provided  $d$  and  $f$  share the same, or compatible, referential bases. If this *metacondition* is not satisfied,  $p$  is not a possibility relative to  $f$ , and in fact has no identity (no *possible* identity) relative to  $f$ . In such a case, relative

to *f*, *d* is, strictly speaking, meaningless.

Referential incoherency will later occupy us in some detail; here, we reach the general conclusion that anything that can be talked about or otherwise referred to without metalogical self-referential inconsistency, without putative reference of a kind that is not permitted in a presupposed framework of reference, is *possible*.

#### 7.4 A general theory of possibility

In looking back over the ten major concepts of possibility we have identified, we may notice two important facts about the nature of possibility. First, we see as we progress along the spectrum of different concepts of possibility a generally decreasing or loosening of “parameters of constraint” (from the Greek: measures of the limits that stand beside). A significant restriction upon what is admitted to be “possible” is found in the *psychologically based* variety, for here what is possible is constrained either by the imaginative, conceivability powers of an individual mind or of a group of minds, or limited by an individual’s particular abilities, to speak, to use a certain language, to employ certain technical skills, etc. These constraints are then loosened in passing from the psychologically based concept of possibility to possibility understood in *temporal* terms, where the possible is limited to what is true, will be true, or perhaps limited to past forgotten events and their consequences that have been exhausted during the passage of time. As we proceed to other concepts of possibility, the constraints that determine *physical* possibility make room for whatever physical laws permit; those that determine *epistemic* possibility allow as possible whatever is not excluded or prohibited by what we now know. The restrictions that govern *formal* possibilities are those that a given formal system establishes, and here the explosion of diverse formal systems shows us that formal possibilities themselves come in great variety. Possibility understood as *complementarity* then loosens classical constraints on possibility.

By the time we arrive at *parametric* possibility, the concept of possibility has broadened still further and has become more purely abstract, general, and comprehensive. *Many worlds* possibility has suggested an open semantic metaphorical conception of possibility, an imaginative/semi-visualizable way of modeling possibilities. *Framework-relative* possibility relinquishes this imaginative appeal in favor of a clear, informative, readily applied general concept of possibility.

Finally, at the most abstract, maximally theoretical general end of the spectrum of possibility lies the *metalogical* variety, which is constrained only

by the preconditions of all identification. Here, we are not restricted to a particular object domain (e.g., psychological or physical events, a particular formal system, etc.), but retain an openness to admit the universe of discourse provided by any identification framework, while, at the same time, affirming the metalogical condition that only metalogically self-referentially consistent descriptions refer to possibilities. This condition constitutes a restriction that applies across frames of reference in general, and is therefore a constraint, a highly general one, upon what is possible.

The ordering I have used in assembling the different varieties of concepts of possibility is not intended to suggest a decrease in the constraints upon what is to be accepted as possible in any kind of incremental, step-by-step way as we move from one concept to the next. But once we have the major concepts of possibility collected before us, we are able to recognize how some concepts are less free and more constrained in terms of what is admissible when compared with others. In this sense, concepts of possibility express varying “parameters of constraint.”

The second fact that we notice in reviewing the major concepts of possibility is that possibility, no matter how it is conceptualized, is always a function of *some* set of constraints; the greater the restrictions these impose, the more limited and specific will be the notion of possibility they determine. As more and more such restrictions are lifted, the resulting concept of possibility becomes increasingly less specific and more inclusive.

It is tempting to generalize further in a quasi-inductive way, and so to think that we might be able entirely to do away with such constraints. But, as we shall see in connection with the concept of horizon, there are certain limits—the necessity of which the metalogic of reference studies—beyond which we cannot go in lifting parameters of constraint: When we trespass beyond those limits, our assertions “short-circuit,” becoming metalogically self-undermining, that is, meaningless. We shall find that some restrictions on what is possible must always be in force (or expressed in a different way: must rationally be enforced) so as to insure the capacity of our frameworks of reference to serve the functions of meaningful reference. Beyond the parameters of constraint that delimit the universe of the possible lies only unintelligibility.

### 7.5 Necessity

So far in this chapter I have focused on the concept of possibility; here it is time to extend that account to its modal sibling, necessity. I have presented a maximally general account of possibility understood in terms of parameters of

constraint, that is, specifically in terms of conditions of framework-relative admissibility. Such constraints may be either “positive” in the sense of being prescriptive, or “negative” in the sense of constituting injunctions that function as prohibitions. Positive constraints, for example those found in chess, are rules that describe moves that must be followed. Negative constraints, for example those embodied in eight of the Ten Commandments, are injunctions that prohibit specified behaviors (“Thou shall *not* ...”).

Taken together, such positive and negative constraints constitute rules that define what is required or prohibited relative to a given frame of reference. As we have seen, as such conditions of admissibility are loosened, the range of what is possible is expanded, and conversely, as conditions of admissibility are tightened, the range of the possible is contracted.

In these terms, possibility has been understood as what is permitted by a system of constraints, so that what is possible is conceptually equivalent to its not being necessary that it not be the case. This equivalence is typically symbolized in the form  $\diamond p \equiv \sim \Box \sim p$ , that is,  $p$  is possible if and only if it is not necessarily the case that not- $p$ . Stated differently but equivalently, what is necessary is conceptually equivalent to its being impossible for it not to be the case, symbolized as  $\Box p \equiv \sim \diamond \sim p$ , that is,  $p$  is necessary if and only if it is impossible that not- $p$ .

Having understood possibility in terms of conditions of framework-relative admissibility, necessity, too, will be understood in a similar way. Where the range of possibility is determined by the looseness of the strictures imposed by a framework’s parameters of constraint, so too do those parameters of constraint determine what must necessarily be the case relative to that framework. What is possible is both what is not prohibited by negative constraints, and what does not conflict with positive constraints. What is necessary is what accords with a framework’s positive constraints while complying with the injunctions imposed by its negative constraints.

The concepts of possibility and necessity are both rule-based and rule-governed framework-relative concepts. Such rules establish framework-relative requirements (necessities) while they also establish boundaries of admissibility (possibilities); the two concepts are conceptually and logically linked with one another. Their linkage will prove to be essential when we study the nature of presuppositions that play a fundamental role on the modal level: presuppositions that are necessary in order for frameworks of reference to be possible.

### 7.6 Excursus: Modal logic and the present study

During the first half of the 1900s, C. I. Lewis and other logicians began to study modal logic by means of a rapidly expanding succession of formalized systems. From a logicist-formalist point of view, this work has been unquestionably fertile: The number of alternative, distinct formal systems of modal logic has exploded, much to the delight of modal logicians, leading to the publication of a large literature. However, for those whose interest lies in a clear-cut semantical understanding of the modal concepts of possibility and necessity and in a detailed grasp of what modal inference means, the proliferation of non-equivalent modal logics has sometimes been felt to be disappointing. Referring to the achievements of logicians in constructing increasingly numerous alternative modal systems, Michael J. Loux has commented: “their work did more to harm than to help the cause of modal logic” (Loux, 1979, p. 16). As he explained:

Lewis and other early figures in twentieth-century modal logic were successful in giving axiomatic presentations of the various modal systems; but while they succeeded in specifying the syntax for those systems, they failed to come up with anything like a thoroughgoing semantics for the various modal systems. They failed, that is, to identify models for those systems, sets of objects in terms of which the formulas of the systems could be interpreted.... [F]or in the absence of a semantics, modal logicians lacked anything more than the ordinary language renderings of  $\Box$  and  $\Diamond$  as ‘Necessarily ...’ and ‘Possibly ...’. A clear-cut semantics for the various systems would provide us with models for the systems, sets of objects we could take the bare formulas of the system to be about; and presumably that would provide us with some clear-cut intuitions as to what a choice from among the various non-equivalent calculi actually involves. (Loux, 1979, p. 19)

While considerable attention has been paid to the formal properties of the now-numerous non-equivalent systems of modal logic,<sup>99</sup> in comparison, much less attention has been given to what the formalized expressions of possibility, necessity, and modal inferences themselves should be interpreted to mean. Just what the preceding phrase ‘should be interpreted to mean’ is itself

---

<sup>99</sup> Hughes and Cresswell (1968, pp. 359-368) provide a detailed formal summary of many of these systems.

intended to mean has been ambiguous: Often philosophers, as exemplified by Loux above, who have wished for more conceptual clarity about modal concepts, have complained that formalized modal systems lack “intuitive meaning,” or that it is unclear how modal formalizations can be associated with models that function as their interpretations, or that there is a lack of applicability of formalized modal logics to actual problems with which such philosophers are concerned.

Perhaps in no other area of philosophical interest have these shortcomings been felt more acutely than in the study of transcendental argumentation, where clarity concerning the central concepts of possibility, necessity, and modal inference is crucial and where some light upon them might be shed by developments in modal logic. In several subsequent sections of this book, readers will encounter periodic use of modal logic; with those future applications in view it will be helpful to include here some specific comments about formalized modal logics.

### *1. Obstacles to the philosophical use of formal systems of modal logic*

There are several reasons why philosophers, and in particular philosophers with an interest in studying transcendental argumentation, have not found formalized modal logics readily useful in their work. For one thing, it must be admitted that many philosophers have not become technically proficient modal logicians, often, I believe, not from lack of potential competence but because they have perceived that applicable formal tools of modal logic have simply not yet been developed which they would need for their specific purposes. In addition, as I view the field, comparatively few philosophers who have cultivated an interest and skills in mathematical logic have found both effective and philosophically informative ways of applying formalized systems of modal logic to their subject-matter. This combination—lack of expertise as logicians, the recognition that modal logic has yet to evolve to a degree that meets many particular philosophical needs, and the fact that even logically proficient philosophers have had only limited success in applying modal mathematical logic effectively and informatively in solving philosophical problems—this combination has served as an obstacle to formalized philosophical work in which modal concepts are central, and especially has this situation hindered the task of formalizing transcendental argumentation.

But beyond these realistic impediments, there are three more specific concerns that are important to mention and which may serve as a preliminary to the discussion that follows. First, because of the now large number of non-equivalent systems of modal logic that have been developed, there is the practical need to select those that give us reason to think that they are appro-

appropriate and applicable to the philosophical subject-matter of the present study; for this, we need one or more justifiable criteria of selection. Second, as we shall see, we encounter immediate difficulties in pinning down the meaning of multiple modal operators, often called ‘higher-order modalities’, such as we find, for example, in the proposition  $\Diamond p \supset \Box \Diamond q$ , that the possibility of  $p$  implies the necessity of the possibility of  $q$ , or in more complex propositions such as  $\Diamond \Diamond p \wedge \Diamond \Diamond q \supset \Box \Diamond (\Diamond p \wedge \Diamond q)$ .<sup>100</sup> Third, a case can be made that non-formalized (and possibly non-formalizable) reasoning may still, despite rapid technical advances in mathematical logic, be more powerful than current formalized tools make possible. This is a question we shall consider later in this study (see Supplement §§2–3).

In this chapter I would like to suggest responses to the first two of these concerns.

## 2. *Choosing a philosophically appropriate system of modal logic*

The expanding universe of distinguishable formalized systems of modal logic can be ordered in terms of the forms of inference that each system authorizes. Given the central interest of this study in transcendental preconditions of reference—in, that is, principles that are necessary in order for reference to be possible—a main criterion in selecting one or more appropriate systems of modal logic will be its or their potential capacity both to express the conceptual content of thought and discourse about such transcendental preconditions and to formalize the reasoning involved in transcendental argumentation.

Early in the history of the development of formalized systems of modal logic the question naturally arose how, or on what basis, are we to judge which system (or systems) is (or are) “correct” or “true,” in order that we might discriminate among the growing multiplicity of formal systems and select one or more best-suited to represent both our nonformalized concepts of possibility and necessity, as well as the logical inferences that we make when using those concepts in nonformalized reasoning.

Logicians who contributed greatly to the birth of formalized modal logic such as C. I. Lewis and Jan Łukasiewicz expressed strong preferences when faced with this choice. Their examples are interesting and can be instructive.

---

<sup>100</sup> Or in relation to a different example: “[T]he problem is not just that we are unsure whether ‘If possibly necessarily  $p$ , then necessarily possibly possibly necessarily  $p$ ’ is a correct modal principle. Few, if any, of us have even the faintest idea what this sentence means. (Loux, 1979, p. 19).



3. *The choice made by C. I. Lewis*

In 1932, C. I. Lewis with co-author C. H. Langford developed a basic formalized system of modal logic known as S1. They accepted ‘ $\diamond$ ’ as an undefined symbol; ‘ $\rightarrow$ ’ as the symbol for strict implication; ‘ $p \rightarrow q$ ’ as equivalent to ‘ $\sim\diamond(p \ \& \ \sim q)$ ’, plus the following basic axioms:

$$\begin{aligned} &(p \ \& \ q) \supset (q \ \& \ p) \\ &(p \ \& \ q) \supset p \\ &p \supset (p \ \& \ p) \\ &((p \ \& \ q) \ \& \ r) \rightarrow (p \ \& \ (q \ \& \ r)) \\ &p \rightarrow \sim \sim p \\ &((p \rightarrow q) \ \& \ (q \rightarrow r)) \rightarrow (p \rightarrow r) \\ &(p \ \& \ (p \rightarrow q)) \rightarrow q \end{aligned}$$

Lewis judged system S1 to be the strictest of the five systems S1 – S5 that he developed; its axioms and theorems are inherited by the other systems of formalized modal logic that he considered to be worth studying. But S1 permits the unrestricted use of so-called ‘higher-order modalities’, such as the necessity of possibility, the necessity of necessity, the possibility of the necessity of necessity, etc. As a result of complications arising from such iterated modal operators, other logicians have advocated adding to S1 such axioms as:<sup>101</sup>

1.  $\sim\diamond\sim p \rightarrow \sim\diamond\sim\sim\diamond\sim p$       (or  $\diamond\diamond p \rightarrow \diamond p$ )
2.  $\diamond p \rightarrow \sim\diamond\sim\diamond p$       (or  $\diamond\sim\diamond p \rightarrow \sim\diamond p$ )

The first of these has been called a ‘weak reduction principle’, and the second a ‘strong reduction principle’. Both function to reduce iterated modalities, but the second does this more than the first.

By applying the equivalence  $\Box p \equiv \sim\diamond\sim p$ , these two principles can be expressed as:

- 1'.  $\Box p \rightarrow \Box\Box p$  and
- 2'.  $\diamond p \rightarrow \Box\diamond p$ .

As noted by Kneale and Kneale (1962, p. 551), adding 1', the weaker reduction principle, to the basic axioms of S1 results in claiming that any necessary proposition is necessarily necessary, while adding the stronger reduction principle, 2', has the consequence that any possible proposition is necessarily

---

<sup>101</sup> Kneale & Kneale (1962, p. 551).

possible, which in effect requires that all modal propositions are necessarily true or necessarily false. If one were to accept 2', propositions involving iterated modalities will collapse into propositions containing only first-order modalities, with the consequence that unique meanings would not be associated with the individual modal operators found in an iterated series. As we shall see later, this consequence is unacceptable in the context of our present purposes.

Lewis rejected both reduction principles and wished to allow for the iteration of higher-order modalities. But—and here this is of particular interest—he admitted that he could provide no rationale or arguments to support this preference. He nonetheless claimed that we would be better served by stricter systems of modal logic that do not allow for the reduction of iterated modal operators.

Prevailing good use in logical inference—the practice in mathematical deduction, for example—is not sufficiently precise and self-conscious to determine clearly which of these five systems [S1 – S5] expresses the acceptable principles of deduction.... Those interested in the merely mathematical properties of such systems of symbolic logic tend to prefer the more comprehensive and less 'strict' systems, such as S5.... The interests of logical study would probably be best served by an exactly opposite tendency. (Lewis, 1932, pp. 501-502]

Why should this be the case? Why would “the interests of logical study” be served best by embracing a system of modal logic that does not provide reduction principles? Why is the preservation of multiple iterated modal operators important and what meaning would they express? As historians of logic William and Mary Kneale asked, “If it is not yet possible to decide whether higher-order modalities can all be reduced to first-order modalities, how shall we ever be able to settle the question? What sort of evidence should we seek and where?” (Kneale & Kneale, 1962, p. 556). The Kneales left the question unanswered.

And yet especially those modal logicians whose interests extend beyond the study of the mathematical properties of modal systems have felt that this question requires an answer. Logician G. H. von Wright, for example, emphasized the central importance of answering it:

Our ‘logical intuitions’, apparently, give no strong indication in favour of any definite answer.... One of the main reasons for this, it seems to me, is the fact that higher-order modal expressions like ‘possibly possible’ or ‘possibly impossible’ have hardly any *use* at all in ordinary or scientific discourse (outside of modal logic). A problem of primary importance, therefore, is to *invent a use* or some kind of ‘equivalent’ of a use for the expressions in question (outside modal logic). (von Wright, 1952, p. 557)

As we shall come to see, to accomplish the objectives of the metalogic of reference there is an important need for higher-order modal expressions. We shall not have to invent a use for them, but will find the need for them to be evident and clear.

#### 4. *The choice made by Łukasiewicz*

Most formalized systems of modal logic that have been developed since the work of C. I. Lewis incorporate rules for the reduction of modal operators, so that, for example, propositions like  $\Diamond\Diamond\Box p$  can be reduced to a simpler form by eliminating multiple prefixing modal operators. As we have seen, the presence or absence of such rules can be pointed to as one way of differentiating and ordering the inflationary universe of modal systems.

Many years ago, Jan Łukasiewicz (1953) proposed a formalized system that he called ‘basic modal logic’. He considered this fundamental system to be “quite inevitable.”<sup>102</sup> Lemmon (1959, p. 48) conveniently summarized the eight conditions that Łukasiewicz’s basic modal logic stipulates:<sup>103</sup>

1.  $\Diamond p \equiv \sim\Box\sim p$
2.  $\Box p \equiv \sim\Diamond\sim p$
3.  $p \supset \Diamond p$
4.  $\Box p \supset p$
5.  $\Diamond p \supset p$  (rejected)
6.  $p \supset \Box p$  (rejected)
7.  $\Diamond p$  (rejected)
8.  $\sim\Box p$  (rejected)

<sup>102</sup> As Henderson (1959, p. 48) expressed this.

<sup>103</sup> Lemmon states these in Polish notation; they are here translated into standard infix notation.

We see that Łukasiewicz accepted the traditional defining equivalences  $\diamond p \equiv \sim \Box \sim p$  and  $\Box p \equiv \sim \Diamond \sim p$ ; he accepted  $p \supset \diamond p$ , but he rejected the reverse implication  $\diamond p \supset p$ ; and he accepted  $\Box p \supset p$ , but he rejected  $p \supset \Box p$ .<sup>104</sup> As we shall see later on, there are good reasons from the standpoint of the metalogic of reference to accept, and to reject, as Łukasiewicz did, precisely the corresponding propositions expressed by the latter two pairs of implications. Łukasiewicz was very strongly committed to these “principles” that defined his basic modal logic, so much so that he claimed: “I call a system ‘modal logic’ if, and only if, it includes the basic modal logic as its part” (Łukasiewicz, 1953, p. 113).<sup>105</sup> Much like Lewis’s system S1, Łukasiewicz’s basic modal logic allows for iterated modal operators, without incorporating reduction rules.

Though he did not deviate from his adherence to these basic modal principles, Łukasiewicz never revealed his rationale that would have explained the strength of his adherence to them.

[H]e seems perfectly content not to question them, but in introducing them he does not say explicitly that they are evident.... [For Łukasiewicz,] basic modal logic purports to be essentially about the world, in some of its most general features. Basic modal logic is not contentless... [it does not concern] only a possible, as opposed to the actual, world” (Henderson, 1959, pp. 49, 56).

This belief—that basic modal logic somehow reflects the nature, the “content,” of the actual world—may have been shared by Lewis, which might explain his intuitive preference for non-reducible higher-order modalities. Whatever reasons might be conjectured to account for the similar intuitive choices embodied in the underlying modal commitments that Lewis and Łukasiewicz expressed, we shall probably never know why they both thought non-reducible higher-order modalities to be significant and important enough that they ought to be retained.

Perhaps we can do somewhat better by offering an explicit rationale for these commitments, one that fits the requirements of the metalogic of reference, a rationale that could at the same time explain the shared intuitions of both Lewis and Łukasiewicz.

---

<sup>104</sup> Again translating Łukasiewicz’s Polish outfix notation into the now more commonly used infix notation.

<sup>105</sup> He later repeated that claim in Łukasiewicz (1957/1951, p. 137).

### 5. *Modal logic appropriate for the metalogic of reference*

It is important that we recognize from the outset that the choice of systems of modal logic that are relevant and applicable to any particular philosophical or other subject-matter is a function of a given set of interests and objectives. Such a choice is an evaluative decision and so rests upon a prior set of values and purposes.

For the purposes of this study of the metalogic of reference we shall find it important to avoid systems that involve modal reduction rules that, with the automatism of a formalized system, collapse multiple modal operators. Certainly, in some cases, we may justify the collapse of multiple or iterated modal operators, but we wish to be free to exercise case-by-case reflective analysis when it is meaningful and appropriate to do this. The brief discussion that follows should help the reader appreciate the need for and the role of higher-order modalities, and subsequent chapters will make this need more evident.

A now well-known system of modal logic is called 'T' (also sometimes called 'M'). It incorporates the propositional calculus equipped with the rule of modus ponens, the rule of necessitation (if  $p$  is an axiom or theorem of T, then  $\Box p$  can also validly be asserted), and the following two modal rules:

- (1)  $\Box(p \supset q) \supset (\Box p \supset \Box q)$
- (2)  $\Box p \supset p$

So far this basic modal logic appears to be transparently understandable in framework-relative terms: (1) tells us that if in a given frame of reference there is a rule in force that establishes the necessity of an if-then relationship between  $p$  and  $q$ , then it will follow that if there is a rule in force that establishes the necessity of  $p$ , then  $q$  will also necessarily be established according to rule. (2) may be understood to state that if in a given frame of reference  $p$  is according to rule necessarily the case, then it will follow that  $p$  is the case.

However, if we attempt to move beyond system T to so-called 'stronger' systems of modal logic, we run into difficulties in understanding in similar framework-relative terms how the rules that define them are to be understood. For instance, Lewis's system S4 is constructed by adding to T the further rule:

- (3)  $\Box p \supset \Box \Box p$ .

This rule of inference, for our present purposes, is objectionable and is rejected, for merely because a framework of reference establishes  $p$  according to rule, it does not follow that there is yet another rule in force according to

which the first rule is set in place. —In other words, when iterated modal operators occur in an expression, we will find it useful in the metalogic of reference to treat each prefixing operator as associated with a different context of discourse, or with a distinct level of abstraction, or with a differentiable framework of reflection. (3) might therefore be more clearly written in the following form:

$$(3') \quad \Box_1 p \supset \Box_2 \Box_1 p.$$

Each prefixing necessity operator is in this way explicitly associated with a distinguishable context of discourse, level of abstraction, or framework of reflective analysis.

In S4, the following two equivalences are derivable:

$$(4) \quad \Box p \equiv \Box \Box p$$

$$(5) \quad \Diamond p \equiv \Diamond \Diamond p.$$

(4) is objectionable for the same reason that (3) above is. However, the right-to-left implication of (4) is acceptable; it states  $\Box_2 \Box_1 p \supset \Box_1 p$ , which, in a framework-relative context, makes good sense: If  $p$  is established according to a first-order rule in a given frame of reference, and that first-order rule is further established by a second-order rule, then it will follow that  $p$  is established according to the first-order rule. (For example, if in chess the game requires<sub>1</sub> that bishops move only diagonally, and if in the palace where chess is played, the sultan requires<sub>2</sub> that the foregoing rule<sub>1</sub> be obeyed (this is rule<sub>2</sub>), then we have this situation: If a bishop is moved diagonally according to rule<sub>1</sub> in the sultan's palace (in accordance with rule<sub>2</sub>), then the bishop will be moved in accordance with rule<sub>1</sub>.)

(5) is objectionable in its left-to-right implication, which we can state in this form:  $\Diamond_1 p \supset \Diamond_2 \Diamond_1 p$ . Simply because  $p$  is permissible<sub>1</sub> in a given framework of reference, it does not follow that its permissibility<sub>1</sub> is rendered permissible<sub>2</sub> by virtue of second-order parameters. (For example, a fishing net permits<sub>1</sub> minnows of a certain size to escape, but it does not follow from this fact that fishing nets of that construction are permitted<sub>2</sub> in a certain country.)

S4 makes it possible to take any expression that makes use of identical iterated modal operators (for example,  $\Diamond \Diamond \Diamond p$ ), and to reduce that string of operators to a single operator (in the example, leaving only  $\Diamond p$ ). This sort of reduction rule for modal operators obscures rather than helps to clarify levels or orders of modality and what they mean; we shall subsequently be concerned precisely with such levels of modality and what they do mean.

In short, a modal system like T offers certain conceptual advantages: It does not authorize the reduction of iterated modal operators, and so when iterated modal operators are called for in the metalogic of reference, each level of iteration is permitted to retain, from a formal standpoint, a non-reducible meaning. Second, a modal system like T may serve our purposes because we shall think of strings of non-reducible iterated modal operators as expressing *levels of contextualization*: Whenever we refer to something as being necessary, or to something as being possible, we must have a context of reference in view. (To be clear, this is not to suggest that such “levels” of contextualization necessarily form logical types: In the expression  $\Box_2\Box_1p$ , the operator  $\Box_2$  is not necessarily to be thought of as of a “higher type” than, or as containing,  $\Box_1$ ; iterated modal operators do not necessarily form nested sequences, although some do. To avoid these connotations, it is preferable simply to regard modal operator subscripts as indicating distinguishable contexts of reference.)

I cannot help but surmise that Łukasiewicz may have had something perhaps similar to our framework-relative reasoning in mind when he insisted that the principles of his basic modal logic were so rock-solid that any acceptable system of modal logic would have to incorporate them. Perhaps Lewis also had something like this in mind when he expressed an intuitive preference for a strict system like S1. As we know, Łukasiewicz’s insistence was not followed by the majority of later modal logicians, and very numerous systems of modal logic have been developed that explicitly repudiate Łukasiewicz’s basic principles. Similarly, later modal logicians have generally not complied with Lewis’s preference to avoid modal reduction rules. Seldom is an effort made to suggest plausible reasons for Łukasiewicz’s or Lewis’s modal choices. Perhaps, we might conjecture, some of those reasons are framework-relative reasons such as we have discussed.

There are a few other ties of Łukasiewicz’s basic modal logic with the logical framework of the present study that bear mentioning before we conclude this section. As noted in an earlier chapter, and as we shall explore further later on, the metalogic of reference requires a logic of more than two truth-values; in this Łukasiewicz was in accord, finding that any system of modal logic must have more than two values.<sup>106</sup> Another affinity with his basic modal logic is that the metalogic of reference also accepts that “true propositions are simply true without being necessary, and false propositions are simply false without being impossible” (Łukasiewicz, 1953, p. 135).

And yet there are significant differences that render the basic modal logic

---

<sup>106</sup> “[I]t is plain that the basic modal logic, and, consequently, every system of modal logic is a many-valued system” (Łukasiewicz, 1957, 113).

of Łukasiewicz unsuitable for our purposes.<sup>107</sup> One of the principal shortcomings of his basic modal logic is that Łukasiewicz does not assign a semantical meaning to his third truth-value (given the number “3”), which in the context of the metalogic of reference would represent the value “meaninglessness.” His truth matrices cannot plausibly be interpreted with this value in mind, as will be made clear later on.

### 7.7 Summary

As we have seen in this chapter, the modal concept of possibility has been understood during the past two millennia in a variety of different ways. In a synoptic overview of this long history, I have identified ten of these varieties, ranked in approximate order of their theoretical abstractness. In this, my purpose has been to show how all of these varieties express the concept of possibility in terms of distinguishable kinds of constraints. As these constraints are loosened and become less and less specifically tied to a subject-matter, we approach at the far end of the spectrum of increasing abstraction a concept of possibility formulated on the level of maximum theoretical generality. I called this maximally general idea ‘the metalogical concept of possibility’.

Possibility understood in these terms is framework-relative in the sense that a frame of reference establishes a range of possible objects of reference; the preconditions of reference that define a given frame of reference establish what is possible from that standpoint. In other words, metalogical possibility is framework-relative possibility, which in turn means identifiable as a function of a particular frame of reference. When we represent the preconditions of reference that form the basis of a given frame of reference, we make explicit what is possible from that standpoint. Such metalogical possibility is constrained only by these preconditions of identification. Such abstract possibility is not tied to any particular object domain; what constrains the range of the possible is the fundamental requirement that only descriptions that are metalogically self-referentially consistent refer to possibilities. This requirement is universal in that it applies across all frames of reference.

Understood in this sense, what is possible is both what is not prohibited by the negative constraints of a frame of reference, and what does not conflict with its positive constraints. Possibility’s sibling, necessity, is then understood in terms of agreement with a framework’s positive constraints, while complying with the injunctions set in place by its negative constraints. From this

---

<sup>107</sup> Rather surprisingly, given his commitment to the fundamental principles of his basic modal logic, its extension to what Łukasiewicz called Ł-modal logic asserts  $\diamond p \supset \diamond p$ , which, for the reasons given earlier, I do not accept. (Łukasiewicz, 1953, Appendix, formula 93, p. 143).



point of view, the concepts of possibility and necessity are rule-based and rule-governed. The rules of constraint that define a reference frame establish, relative to that framework, requirements—that is, necessities—that must be satisfied by object identifications in terms of that framework, while, at the same time, those rules set in place boundaries of admissibility—that is, a range of possibilities of reference.

It is commonly thought that the most logically systematic analysis of the modal concepts of possibility and necessity is to be found through the resources of formalized modal logic. I therefore devoted a portion of this chapter to respond to the two questions, How or on what basis can we make a justified selection of one or more systems of modal logic from the now well-populated universe of non-equivalent systems?, and the question, How are higher-order modalities to be understood? Because of their affinities with the modal logic applicable to the present study, I discussed formalized modal systems developed by C. I. Lewis and Jan Łukasiewicz, with a specific interest in their generally shared commitment to systems of modal logic that preserve iterated modal operators. The view that emerged from this discussion is sympathetic with the rejection of certain forms of modal inference, such as  $\Diamond p \supset p$  and  $p \supset \Box p$ , while accepting the validity of inferences such as  $p \supset \Diamond p$  and  $\Box p \supset p$ . With respect to higher-order modalities, we shall regard these as expressing levels of contextualization, so that when a modal operator refers to a possibility or necessity, a corresponding context of reference is to be understood.

With these as stepping stones, we shall later be able to develop a clear conception of the ways in which a framework-relative understanding of the concepts of possibility and necessity tie together in an analysis of the transcendental preconditions of all referring.

## 8

### Presupposition Logic, Reference, and Identification

“To stand within it is to be unable to see it.”

– Said of the massive volcanic caldera,  
the Campi Flegrei, on which much  
of Naples has been built

[W]e are interested most of all in the presuppositions from which each philosopher draws his dialectic reserves. These often determine the result of all his scientific labours. The most fundamental and profound presuppositions may seem quite unimportant to the man himself, because he takes them as self-evident. Indeed one often discovers, when looking for presuppositions, that the thinker in question did not even know he was presupposing them. He just took them over unconsciously from the general circle of ideas common to people of his time. It is our business to uncover them, and to see what part they played, unbeknown to him, in the progress of his enquiries.

– Leonard Nelson (1970, p. 15)

A frequent complaint among philosophers when thinking critically about the thought of non-philosophers is that the most basic concepts and presuppositions employed by the latter are often passed over in silence, neglected, and ignored. That complaint expresses one of the fundamental reasons for engaging in philosophical inquiry.

The fact that all of us, including even those philosophers just mentioned, make habitual use of basic concepts and presuppositions, so much so that our very mental activity remains unconscious of them, results in our facing an intellectually difficult and challenging task should we wish to render those

concepts and presuppositions explicit and then evaluate them reflectively and dispassionately. Since our most fundamental concepts and presuppositions comprise the very foundation of our conceptual activity, they resist all but the most determined efforts to make them explicit and to subject them to critical appraisal. A main reason why this is so is that, in stepping back reflectively and in seeking to engage in dispassionate self-evaluation, we often bring along with us the very concepts and presuppositions that we would place clearly in reflective focus. If in our efforts to reach critically reflective philosophical understanding we inadvertently carry along the very substructure of our thinking that we wish to stand back from and assess, we cannot get very far.

Our challenge is therefore to find ways of stepping back or out of our habitual reference frames, and if necessary develop new approaches that make it possible to render basic concepts and presuppositions explicit enough so that they can be thought about without being themselves used in the process without our knowing it. This is, of course, much easier to say than it is to accomplish.

In this chapter, I look closely at the logic of presuppositions, and then relate this logic to the concepts of reference and identification that are central to the metalogic we shall develop.

### 8.1 The priority of presuppositions

The ‘pre’ in the word ‘presupposition’ conveys priority, but priority of what kind? Is a presupposition a statement that must be granted prior to other statements that are somehow based on it in order for those other statements to be valid or true? Or is a presupposition a condition, not a statement, that must be in place, hold, or be satisfied prior to other consequences that may follow as a result? Is the “priority” in question a kind of temporal priority, or is it purely logical, or is it in some sense structural?

A wide variety of answers has been proposed to questions like these. Some of these answers are instructive and merit our attention; some varieties can overlap others; some have been studied in the literature, others have not.

#### 1. *Presuppositions as preliminary assumptions*

We often regard the preliminary assumptions that we make before undertaking an action or advancing an argument to comprise presuppositions. In this sense, we commonly believe that the criminal justice system would not be necessary were there not people who break laws, or that Noah’s need for open-heart surgery has been recognized as a consequence of prior preliminary

tests that show occlusion of one of his coronary arteries, or that rallying society to combat random terrorist killings is based on the prior assumption that a country's citizens are owed the protection and safety which their governments can provide. Preliminary assumptions of this kind are frequent, common, and among the most easily acknowledged presuppositions.

## 2. *Conditional presuppositions*

We often set down provisos for subsequent actions, decisions, candidacy, qualification, etc. Such provisos function as prerequisites that must be met in order for their consequences to be realized. For example, completion of a certain graduate degree may in this sense be presupposed as a condition in order to obtain certification to practice a particular profession. The terms of contracts, legislation, civil and criminal law are laden with such familiar conditional presuppositions.

## 3. *Presuppositions as logical premises*

Often the conclusions to which we assent are based on explicitly stated or implicitly assumed premises that form a part of a logical argument. Once those premises have been made clear, we are in a position to decide whether our conclusions follow validly from them. Presuppositions in this sense are hypotheses or assumptions that can be formulated and placed within the framework of a proposed logical argument or deductive proof. Presuppositions in the sense of logical premises comprise a usually clear, unambiguous variety.

## 4. *Presuppositions as logically antecedent suppositions*

Related to the previous two varieties, we often make suppositions that are logically antecedent to an argument, position, or action that we want to advance or propose. Such *pre-suppositions* are seldom explicitly formulated as are logical premises; they are usually "background suppositions" that we take for granted and simply assume without directing attention to them. If a hurricane has destroyed a bridge that is essential to transporting food and water to people on the other side, when a step-by-step logically formulated plan is advanced to rebuild or replace the bridge, such a plan presupposes an antecedent and implicit agreement concerning the value of supplying food and water to the population in need.

## 5. *Presuppositions as the ingredients of definitions*

Not all of those instances for which we regard something to comprise a presupposition are temporally prior facts or presumptions, or can be inserted into

a deductive logical schema. Some are the ingredients of definitions. When we assert that Ted is Jim's uncle, we presuppose as an ingredient of definition that Jim has a father who has a brother named 'Ted'. Such presuppositions are not logical implications, but rather they comprise the defining ingredients of concepts—here, the concept of uncle.

#### 6. *Presuppositions as suppositions of language use*

A distinction is often drawn between, on the one hand, presupposing *statements* or *conditions* that must be granted in order for a given statement to be either true or false, or for a question, injunction, or petition to communicate its intention, and, on the other hand, the *beliefs* that speakers and/or hearers must hold in order for them to regard that statement to be either true or false. I will refer to this distinction in terms of *semantical presuppositions* in contrast to *psychological presuppositions*.

The following are examples of psychological presuppositions: the belief that there is such an object of reference as “black matter” and the belief that “black matter” exists in the physical universe, beliefs that are psychologically implicit when a speaker claims that the statement “Black matter comprises a large proportion of the total matter in the universe” is true. When psychological assumptions of this kind are in view, we say that it is not “correct” or “appropriate” or “reasonable” to make such a statement unless the speaker holds the relevant beliefs—that he actually *believes* what his statement psychologically assumes. Such psychological belief-based presuppositions that are associated with a sentence or statement are conditions which a speaker ordinarily expects establish a common ground of communication between speaker and hearer when that sentence is uttered.

When we use language to communicate what we wish to assert, we routinely do presuppose that our listeners will understand what we intend by virtue of facts or conditions that are implicit in what we have said, even though those facts or conditions are not explicitly stated. Such “speaker presuppositions” are made explicit when we identify what is suggested or intended but not actually stated by a speaker. If I say to a listener that my aunt has red hair, the listener will routinely and conventionally understand that I am implying I have an aunt. Such belief-based “contextual implication” has been the focus of much recent work relating to linguistic presuppositions.<sup>108</sup> Such work reflects the view that the proper object of the philosophical study of presuppo-

---

<sup>108</sup> Sample publications that reflect this approach include: Abbott (2000, 2006, 2008), Beaver (1992, 1997, 2001, 2008), Beaver & Kraemer (2001), Langendoen & Savin (1971), Schlenker (2007, 2008, 2009), Sellars (1954), Simons (2001, 2003, 2004, 2006, 2007), Stalnaker (1972, 1973, 1974, 1998).

sitions is not especially what words, sentences, statements, concepts, and theories presuppose, but what people actually presuppose when they are speaking.

Psychological belief presumptions of this sort will not, however, concern us in this study of the logic of presuppositions.<sup>109</sup> Excluding them from consideration does not mean they are unimportant, only that they are not, strictly speaking, of interest from the standpoint here of the logic of presupposition. Psychological assumptions of the type in question are often a matter of established conventions, of habits of belief that we have developed, or of expectations we have come to form as a result of hearing others make statements, ask questions, utter commands, etc. Social, cultural, anthropological linguistic conventions, habits, and expectations are not our chosen subject-matter.

### *7. Presuppositions of pragmatical activity*

Many of our statements about others or about ourselves relate to activities we engage in. Often we make statements such as: “James did poorly on his bar exam.” This statement presupposes that an activity was in fact engaged in, namely that James actually *took* the bar exam. If he had the flu and couldn’t take the exam, it makes no sense to say that he did poorly on it. The activity presupposed by the statement is the activity of taking the exam. Similarly, we might interpret Descartes’ *Cogito* argument as an attempt to point to a presupposition of pragmatical activity that is unavoidably made when Descartes became self-aware of his own thinking, or when he expressed that self-awareness in the statement “I think.”

### *8. Presuppositions of missing premises*

Not all arguments are complete as they stand. Some, in order to qualify as valid, require supplementation with premises that have not been stated or have been overlooked. Presuppositions of this kind are what an incomplete argument requires in order to be made complete and valid. Philosophers, whose profession it is to develop valid arguments in support of their positions,

---

<sup>109</sup> In this connection, it is important to note that studies of the belief-based linguistic presuppositions of ordinary speech and conversation frequently employ a concept of “projection” that is distinct and unrelated to the meaning of that concept in the present work. The linguistic notion of “projection” relates to the view that a statement or utterance “projects” certain assumptions, its linguistic-pragmatical “presuppositions.” E.g., “The man dressed in scarlet punched the woman” is said in this sense to project “There is a man dressed in scarlet,” “There is a woman who was punched,” etc. The linguistic use of the term ‘projection’ began with Langendoen & Savin (1971); the development of the metalogical concept of projection is first found in Bartlett (1970).

sometimes of course fall victims to the shortcoming of advancing incompletely developed arguments.<sup>110</sup> As a result, a major way in which a philosophical argument may be criticized is to show that its stated premises are inadequate to prove its conclusion. It can be important to understand just what such criticism actually means, even when we are in a forgiving mood:

When we say Smith's argument requires P, do we mean the argument Smith stated or the one he meant to state? There is some danger of confusion here. The argument he stated, A, was invalid, and that invalid argument cannot be "made valid" by adding extra premisses. It is what it is, an argument with insufficient premisses. Brown notes that there is another argument B, consisting of A + P, and that B is valid. Brown may charitably decide that B was the argument Smith "really had in mind." In which case, when Brown says that argument A assumes or presupposes P, we may charitably suppose he really means:

Argument A is invalid

Argument B = (A + P) is valid. (Palmer, 1985, p. 100)

And so, if we are in such a charitable state of mind, we reinterpret Smith's original argument so as to give him the benefit of his missing premisses, and generously allow that he was really "presupposing" argument B. Charity or not, premisses of an argument that are presupposed but that are unrecognized or left unstated comprise a distinguishable variety of presupposition.

### 9. *Presuppositions of existence*

A very considerable amount of time and effort by dedicated philosophers has been expended, some might say squandered, in connection with presuppositions of existence. Suppose someone states, "The tree in Indonesia that was 7,000 years old has recently been cut down." It is natural and reasonable to recognize the underlying presupposition of existence, "There was such a tree." If there never was such a tree, the preceding statement is of course not true, but then it is not false either, if its falsity is equated with the truth of the statement "The tree in Indonesia that was 7,000 years old was *not* recently cut down." If there never was such an old tree in Indonesia, the original statement as well as its negation fail to refer. For a statement to be true or false, it must

---

<sup>110</sup> This frequently overlooked variety of presupposition has been studied in detail by Palmer (1985, Chap. 7, §37).

somehow refer.

Whether and under what conditions statements, questions, instructions, commands, and so on rest on presuppositions of existence, is, as philosophers are wont to say, “a thorny issue.” We should be able to de-barb those thorns later on by means of some explicit attention to framework relativity.

#### 10. *Linguistic presuppositions of reference*

Two previously mentioned varieties of presupposition—presuppositions of pragmatical activity and presuppositions of existence—are sometimes lumped together as belonging to the class of linguistic presuppositions of reference. Philosophers who bundle them together do this because both kinds of presupposition can be understood in the context of analyzing language use in truth-functional terms, in the following way: Again here is our example of a presupposition of pragmatical activity:

Statement *S*:            James did poorly on his bar exam.  
 Presupposition *P*:      James took the exam.  
 If presupposition *P* is not met, *S* is neither true nor false.

And again the example of a presupposition of existence:

Statement *S*:            The tree in Indonesia that was 7,000  
                                       years old has recently been cut down.  
 Presupposition *P*:      There was such a tree in Indonesia.  
 If presupposition *P* is not met, *S* is neither true nor false.

In both examples, *S* presupposes *P*, and when *P* is not satisfied, *S* may rightly be considered to be neither true nor false. (When *P* is not satisfied we might instead prefer to consider *S* to be “pointless,” “odd,” “inappropriate,” or even, under some circumstances, “meaningless.” It is not my intent here to collect reasons why it is better to choose one of these alternatives rather than the others.)

The above truth-functional understanding of presuppositions of pragmatical activity and of presuppositions of existence has been challenged, criticized, defended, revised, or replaced by its objectors or propounders. We look more closely at the logic involved in the truth-functional understanding of such presuppositions later in this chapter.



### 11. *Presuppositions of concepts*

When we turn to consider the presuppositions of concepts, we no longer are primarily interested in analyzing the presuppositions of linguistic expression, but rather the conditions that are basic to the meaning of whatever concepts we have in view. For example, suppose we consider skepticism abstractly as a theoretical position, and not as a particular linguistic expression uttered by an individual, and suppose that according to that position, there are no truths. The position itself makes recourse to the concept of truth, presumably in its publically understood, familiar meaning. Skepticism that claims there are no truths is a position that sets itself in conflict with what it must presuppose in order for the concept of truth to have the meaning that skepticism intends.

Or to take another example: Lying is a meaningful concept. If it is, telling the truth must also be a meaningful concept, for the meaning of the concept of lying presupposes that of truth-telling—the notion of lying, which is not to tell the truth, would have no meaning otherwise.

### 12. *Structural and systemic presuppositions*

Presuppositions of this sort are conditions without which an object, organized collection, or interrelated system loses its capacity to cohere or function as an integrated ensemble. The cables of a suspension bridge serve as structural presuppositions in this sense, as may a line, or group of lines, of code in a computer program. Such structural and systemic presuppositions are commonplace, for example, in engineering and the theory of general systems, and they are central, as we shall see later on, to the metalogic of reference.

### 13. *Presuppositions of identification*

These are conditions without which the identity of an object of reference cannot be recognized. Of the varieties of presupposition so far described, these are theoretically the most fundamental, general, and ubiquitous in no matter what area of human thought or discourse we may wish to consider. We'll return to consider presuppositions of identification in more detail in a moment.

...

To summarize, the following are varieties of presupposition that we have so far recognized:

1. presuppositions as preliminary assumptions
2. conditional presuppositions

3. presuppositions as logical premises
4. presuppositions as logically antecedent suppositions
5. presuppositions as the ingredients of definitions
6. presuppositions as suppositions of language use
7. presuppositions of pragmatical activity
8. presuppositions of missing premises
9. presuppositions of existence
10. linguistic presuppositions of reference
11. presuppositions of concepts
12. structural and systemic presuppositions
13. presuppositions of identification

There are, to be sure, sufficient commonalities among these 13 varieties of presupposition to justify grouping them together. For in each, and in diverse ways, a relation is pointed to that purports to consist in or to establish some form of “necessity” that is thought to hold between a presupposer and that which it presupposes. What is presupposed “must,” in some sense, be granted, allowed, or be implicitly accepted in order for that which presupposes it to happen, obtain, be the case, be true, be meaningful, or be identified.

However, there are important differences among these varieties of presupposition, differences that will prove to be, in the context of the present study, of more importance than what they have in common. It is important that we recognize certain of these differences, which we’ll do in the next section.

The observant reader will notice that, somewhat as we saw in the last chapter in connection with varieties of possibility, the varieties of presupposition I have identified become, in a general and not strictly step-by-step way, increasingly inclusive. That is, the later varieties of presupposition are theoretically more abstract and have an increasingly greater range of applicability. Since the focus of interest in this study is a certain group of concepts, relations, and structures of maximum theoretical generality, the later, more highly general varieties of presupposition possess for us the most interest. The varieties of presupposition that will particularly concern us are, in fact, the last five (9. –13.), and some of these more than others.

We turn now to consider the logic underlying these forms of presupposition. To do this, imagine that we “slice” across the latter five classes of presupposition in a manner so that each “slice” is associated with one of three general ways in which distinguishable varieties of presupposition can be ex-

amined: as a function of truth, as a function of meaning, or as a function of identifiability. We discuss the first and last in this chapter, only touching on the second, which will be reserved for more thorough study in later chapters.

## 8.2 Truth-functional presuppositions

Presuppositions that lend themselves to truth-functional analysis are presuppositions that seem to be inextricably carried along by or somehow built into referring statements (when appropriate, here, as elsewhere, I include referring questions, commands, prohibitions, etc.). I shall use the term ‘truth-functional’ in a broad sense in which attempts are made to understand sentences, statements, and propositions as a function of the truth or falsity of their presuppositions (alternatively, whether those presuppositions are satisfied or are met). ‘Truth-functional’ in this wide sense will not preclude that such sentences, statements, and propositions may sometimes *not* possess the values “truth” or “falsity,” but may instead have other values, such as “meaningless,” “inappropriate,” “odd,” etc. When this happens in a “truth-functional context” it is because there is a failure of presuppositions to support the truth or falsity of statements that presuppose them. In a truth-functional context, as we shall see, the truth or falsity of such statements is functionally tied to relevant presuppositions that must hold.

We have already encountered presuppositions that are referential in a truth-functional sense under the headings of presuppositions of existence and of linguistic presuppositions of reference. A now long-standing tradition has been established that claims that the logic of such presuppositions can best be understood as a function of the truth or falsity of referring statements and of their associated presuppositions. In this tradition, Strawson’s views<sup>111</sup> continue to remain central in philosophical discussions.

The most widely accepted logic of truth-functional presuppositions may perhaps most clearly be described by clarifying the nature of three kinds of logical relation: implication, entailment, and referential presuppositions of truth and falsity.<sup>112</sup>

---

<sup>111</sup> Strawson (1970/1950, 1952, 1970a/1954). His views over time underwent some change. Since exposition is not the purpose here, reference is made here only to his earlier position.

<sup>112</sup> For readers interested in consulting some of the earliest 20th century works that helped to set the basis for the study of these varieties of logical relation, see, for example, Hancock (1960), Nerlich (1965), van Fraassen (1968), Ginsberg (1972), and Kempson (1975).

1. *Necessary and sufficient conditions of truth: logical implication*

Using ‘ $\supset$ ’ to symbolize logical implication, also called ‘material implication’, we recognize that A is a *sufficient* condition for B if when A is true, B must be true—i.e.,  $A \supset B$ . “If it rains, the humidity goes up”—raining then is a sufficient condition for rising humidity. A is a *necessary* condition for B, or A is *indispensable* for B to be the case, i.e., only if A then B—so if B is true, then A must be true since A is indispensable for B to be true. That is,  $B \supset A$ . In the Sahara Desert, it might be the case that “Only if it rains, will the humidity go up”; raining would then be a necessary condition for rising humidity.

Let’s consider a more detailed example: Suppose there is an electrical circuit in which power is supplied by a battery, and a switch controls the flow of current from the battery to a light bulb. We assume that the light bulb will not light up unless power from the battery reaches it. Electrical power is then indispensable to the bulb lighting; electrical current is a necessary condition for the bulb to light up.

Let us call the switch ‘A’ and the bulb ‘B’. Only if switch A is in the “on” position ( $A = \text{true}$ ) will bulb B light up ( $B = \text{true}$ ). In other words, only if A then B. If bulb B lights up, switch A must be in the “on” position: Switch A being in the “on” position is indispensable for bulb B to light up. Hence in this analogy, A must be true for B to be true. A is a necessary condition for B. But A is not a sufficient condition for B: It is possible that switch A is thrown (is placed in the “on” position, or  $A = \text{true}$ ), but bulb B may not light up. This would happen if the battery that supplies power to the circuit is dead or is disconnected. This possibility illustrates that satisfying a necessary condition (here throwing switch A) does not thereby automatically amount to satisfying a sufficient condition (to turn on bulb B). But if light bulb B *does* light up, we do know that switch A must be in the “on” position; A is necessary for B.

The following truth-table for implication makes these alternatives clear:

“The implication relation holds” means that electrical current flows, or does not flow, in accordance with the circuit described above: only if the switch A is on does the bulb B light up

B: the bulb, can be either on or off

A: the position of the switch, on or off

<b>B</b>	<b>A</b>	<b><math>B \supset A</math></b>	
on	on	T	Here the implication relation holds: when B is on, A is on
on	off	F	This alternative is ruled out: for the implication relation to hold, B cannot be on and A off
off	on	T	The implication relation holds: A may be on even though B is off (e.g., the battery is dead)
off	off	T	Again the implication relation holds: both A and B may be off

## 2. *Entailment*

There is a second important and related variety of logical relation that we need to understand clearly, that of *entailment*. We'll symbolize the entailment relation by means of the derivability sign ' $\vdash$ '. A relation of entailment describes the connection between one or more statements whose truth logically necessitates the truth of one or more other statements, that is, whose truth follows necessarily from them. Entailment relations are always embedded, implicitly or explicitly, in a context. The context of a relation of formal logical entailment would be a formalized system of logic in terms of which one or more statements whose truth can be derived by following the rules of deduction authorized by that system. For example, in standard propositional calculus, from P and  $P \supset Q$ , Q can be derived by the rule of modus ponens, i.e.:  $P, (P \supset Q) \vdash Q$ . Q can be derived, its truth can be proved, on the basis of the truth of the statements to the left of the derivability sign, by employing one or more rules of deduction authorized by the logical system that provides the context for the proof.

The following example relies on a different context of entailment, one that is a *function of meaning*: If "Joe is a man" (A) is true, then "Joe is a human being" (B) must be true. The truth of A entails the truth of B. Here, the context is determined by the meaning of 'man' which includes that of 'human being'. The rules of derivation of a formal system of logic are not explicitly applied (though in our informal reasoning we may be presupposing them); rather, the definitions of key terms or concepts play the main role and estab-

lish the context of the entailment relation between A and B, a relation which is a function of the meanings of A and B.

Entailment, then, is a relation between statements, sentences, or propositions which is such that if, in a certain context, A entails B, B follows necessarily from A. The entailment relation may express the relation of derivability in a formal system of logic, or it may express a necessary relation between the meanings of propositions, one whose meaning requires that the other also hold.

### 3. *Referential presuppositions of truth and falsity*

Let us look at a different example, this time involving what is explicitly a truth-functional referring presupposition: A married couple sends their back door out to be refinished by a local woodworking company. That night, while the back door is not in its accustomed door-frame and is still in the woodshop, the husband, out of habit, absent-mindedly asks his wife, "Is the back door locked?" (We'll call this 'question B'.) They both laugh: The husband has said something foolish, for question B presupposes the truth of the referring statement A "The back door is in place," or "This house has an existing back door," which at the time of the husband's question is not the case. Note that statement A is just as much presupposed by the contrary question "Is the back door unlocked?" Some philosophers have claimed that since the presupposition A is not satisfied (or is false), the question raised by the husband is without meaning, or is inappropriate, or is at least odd, or laughable. It will not be important for our purposes here to choose among these interpretations. Suffice it to say that statements or questions like B, which presuppose that another statement A is true, are considered to be neither true nor false when A is not satisfied or is false.

The logic of truth-functional referring presuppositions differs then in a fundamental way from that of both implication and entailment: If statement B is in view and we wish to identify what B presupposes, we are then interested in specifying those statements that must be true in order for B to have *either* truth-value, either true *or* false.

We cannot construct a parallel two-valued table for the truth-functional presupposition relation, since, in the now-standard conception, its logic must be three-valued: If B presupposes A, and A is not the case, B is neither true nor false. We might call its value 'indeterminate' (I)—or inappropriate, foolish, truthvalueless, etc. Asking the question "Is the back door locked" presupposes that the back door is there; when it's not there, the question no longer is appropriate.

Let ' $\wp$ ' symbolize the three-valued referential relation ' $\dots$  presupposes  $\dots$ '. We can construct a table that summarizes *just* when the relation  $\wp$  holds; the presupposition relation  $\wp$  holds in the following cases (note that there are only three alternatives now since when presupposition A is not met, B is neither true nor false, but indeterminate, I):

<b>B</b>	<b>A</b>	<b>B <math>\wp</math> A</b>	
T	T	T	When B is true, A is true
I	F	T	A, which is presupposed by B, is false; B is therefore neither true nor false, but indeterminate (I)
F	T	T	B presupposes the truth of A even when B is false

In short, the presupposition relation  $\wp$  holds only when (1) if B is true, A is true, (2) if A is false or is not satisfied, B is neither true nor false (and if B is a question, command, etc., it is inappropriate, senseless, etc.), and (3) if B is false, A must be true or be satisfied.

To make the contrast between presupposition and entailment explicit, here is a table that summarizes just when the entailment relation  $\vdash$  holds:

<b>B</b>	<b>A</b>	<b>B <math>\vdash</math> A</b>	
T	T	T	When B is true, A's truth is derivable from B

For the entailment relation  $\vdash$  to hold between B and A, when B is true, A must be, so that if A is true, and if  $B \vdash A$ , then B could not be false since A's truth is derivable from B.

Note that the following case cannot arise, provided that A is derivable from B:

<b>B</b>	<b>A</b>	<b>B <math>\vdash</math> A</b>	
T	F	F	If $B \vdash A$ were true, then when B is true, A's truth would be derivable from B, so since A here is false, $B \vdash A$ is cannot be true

And then there are the two remaining cases:

<b>B</b>	<b>A</b>	<b>B ⊢ A</b>
F	T	
F	F	

When B is false, the derivability of A does not arise—no proof of A from B is possible in the entailment context that at issue; suppose we call this system ‘S’. Since the derivability of A doesn’t arise when B is false, we leave two blanks in the above table.<sup>113</sup> But, independently of context S, A might still be true (or false), but of course the derivability, *in S*, of A from B when B is true is unaffected.

In short, we shall understand the entailment relation in the following sense: Entailment in an appropriate entailment context S holds only when (1) if B is true, A must be true, (2) if A is false, B must be false, and (3) if B is false, A’s truth is undecided.

In comparison with entailment, we recall that the truth-functional referring presupposition relation  $\wp$  in the statement “B  $\wp$  A” holds only when (1) if B is true, A is true, (2) if A is false or is not satisfied, B is neither true nor false (and if B is a question, command, etc., B is inappropriate, senseless, etc.), and (3) if B is false, A must be true or be satisfied.

The table below summarizes some of the main observations we have made:

<b>Implication</b>			<b>Entailment</b>			<b>Presupposition</b>		
<b>B</b>	<b>A</b>	<b>B ⊃ A</b>	<b>B</b>	<b>A</b>	<b>B ⊢ A</b>	<b>B</b>	<b>A</b>	<b>B <math>\wp</math> A</b>
T	T	T	T	T	T	T	T	T
T	F	F	T	F	F	I	F	T
F	T	T	F	T		F	T	T
F	F	T	F	F				

<sup>113</sup> Many logicians believe that if B here is false, this constitutes a sufficient condition for A to be false; they therefore reason that the two blank cells in the above table should be filled with F’s: The entailment relation does not hold when B is false, so they consider the entailment relation to be false. (See, for example, Kempson (1975, p. 48).) Others, however, have argued that when B is false, it may happen that B has presuppositions that A does not have, and so then it is possible for B to be without a truth-value (i.e., to be neither true nor false) when that presupposition is not met, while at the same time A could be either true or false. Complexities like these result when entailments themselves involve presuppositions. For purposes here, we do not need to consider such situations, and for clarity exclude them. (See, for example, Ginsberg, 1972, pp. 512ff.)



We have now defined three forms of truth-functional relation, each one of which expresses a different concept of truth-functional *necessity*. In the case of logical, or material, implication, the necessity that logical implication expresses is that of the truth of one statement that serves as a necessary condition for the truth of another. The two statements need have no connection with one another in terms of meaning. For example consider the implication: only if “It rains today in Belgium” (A), then “I will eat oatmeal for breakfast” (B)—there is no detectable relation of meaning between rain falling in Belgium and having oatmeal for breakfast. Even so, the implication lends itself to the same truth-functional understanding as does an implication where there is a meaningful connection: Only if “ $1 + 3 = 4$ ,” then “4 is the sum of the first two odd integers.”

We notice two things: The variety of necessity involved in logical implication is, first of all, indifferent to meaning, and, secondly, the necessity involved is merely stated, but is not proved deductively: No logical rules of derivation justify a logically necessary inference from the truth B (having oatmeal of breakfast) to the truth of A (raining in Belgium).

The relation of entailment is significantly different:

(i) When two statements, or two groups of statements, are connected by a relation of entailment, the truth of the second is logically derivable from the truth of the first. In the case of entailment, the necessary connection between two sets of statements may be a matter of logical derivation, where the second is actually demonstrated logically to follow from the first. Then the concept of necessity that is involved relates to the fact that one set of statements follows from the other set in accordance with the rules of derivation defining the system of logic that is being employed. A proof of an entailment relation can proceed effectively without attaching any interpretation to the symbols employed; such a proof is then purely formal.

(ii) Alternatively, when the entailment context involves the definitions of terms or of concepts (e.g., “Joe is a man” entails “Joe is a human being”) the second is derivable from the first based on the meanings of the terms or concepts involved. Here entailment relates to the variety of presupposition called ‘presuppositions of concepts’, and is a function of their meaning.

The variety of necessity involved in the entailment relation is therefore a direct expression of the rules that determine the *context of entailment*—in the two contexts I have mentioned, these are logical rules of derivation or rules establishing the meaning/use of such expressions as ‘man’ and ‘human being’.

The relation that is involved in presuppositions that are a function of the truth or falsity of referring statements is more complex: Often, it is not immediately evident that the truth of one statement, or its falsity, presupposes the

truth of another statement. The “back door” example given earlier is simple, but when philosophers begin to analyze the presuppositions of more complex statements, often embodied in and integrated into a total philosophical position, the identification and justification of presuppositions can be far from simple.

This short logical analysis of three forms of truth-functional relation enables us to recognize that, in the case of truth-functional referential presuppositions, unlike instances of logical implication or entailment, we lack a method by means of which to test whether, or to justify when, an alleged presupposition really is presupposed by a given statement or group of statements. When we deal with implications or entailments, it is most often possible to justify very explicitly and systematically the claims that we make about them. This is much less so when referential truth-functional presuppositions are studied. The lack of a method to test presuppositions becomes even more significant when we deal with less commonly discussed varieties of presupposition. Later in connection with certain of these special varieties, we shall devise specific tests to meet this need.

### 8.3 Structural and systemic presuppositions

Presupposition seems a purely linguistic relationship, connecting actual statements, concepts or arguments to the persons, ingredients or premisses they do or should contain. But these linguistic relations are remarked on because they reflect underlying non-linguistic facts.

– Humphrey Palmer (1985, p. 103)

In the previous section of this chapter, the presuppositions we have considered have been truth-functional presuppositions that are language-based. The language may be a formalized language, as in many instances that involve material implication or logical entailment, or the language may be a natural language—the ordinary language of everyday discourse or technical languages as used in the various disciplines. But not all presuppositions are language-based.

We are reminded that in the general presupposition theory developed here, a presupposition is any one of various forms of relation that purports to consist in or to establish some form of “necessity” thought to hold between a presupposer and that which it presupposes. What is presupposed needs to be, in some sense, granted, allowed, or implicitly accepted in order for that which

presupposes it to happen, obtain, be the case, be true, be meaningful, or be identified.

Understood in this broad-spectrum sense, it will be evident that not all presuppositions are truth-functional and linguistic in nature. In this section we turn to look at two intertwined varieties of presupposition that have not been studied in the literature, those that are *structural* and *systemic*. We recall from the earlier listing in this chapter that presuppositions of this sort are conditions upon which an object, an organized collection, or an interrelated system necessarily depends in order to cohere or function as an integrated ensemble, and without which that coherence and integration are lost.

Earlier I mentioned the example of the cables of a suspension bridge which serve as structural presuppositions in this sense, as may a line or group of lines of code in a computer program. Such structural and systemic presuppositions are commonplace, for example, in engineering and the theory of general systems, and they will prove to be central, as we shall see later on, to the metalogic of reference.

For the purposes of simplifying the discussion, I separate structural and systemic presuppositions, but we'll see that each variety involves the other: they are, in fact, two sides of the same coin.

### 1. *Structural presuppositions*

The structural presuppositions that I wish to consider are of a theoretically more general nature than the earlier concrete physical example of suspension bridges and the cables that hold them up. In Bartlett (1970), I developed the concept of "*logic of structure*," which bears directly on the nature of structural presuppositions: "The logical, or pure theoretic relational, scaffolding of the theoretic content of a system constitutes that system's logic of structure" (Bartlett, 1970, p. 1). In that work, by the term 'structure' I had in mind Hjelmlev's meaning of "*une entité autonome de dépendances internes*"<sup>114</sup> [an independent entity consisting of internal dependences], "*un tout formé de phénomènes solidaires, tels que chacun dépend des autres et ne peut être ce qu'il est que dans et par sa relation avec eux*"<sup>115</sup> [a whole made up of substantive phenomena, such that each depends on the others and cannot be what it is except in its relation to those others]. (As we shall see in the next sub-section, much the same understanding is to be found in general system theory's concept of "system," hence the intersection and mutual involvement of the concepts of structure and system.)

---

<sup>114</sup> Hjelmlev (1947, p. 69).

<sup>115</sup> Hjelmlev (1953, pp. 638f).

In the present study, I shall use the perhaps clearer phrase ‘*constitutive structure*’ instead of ‘logic of structure’, with both phrases possessing the same meaning. In Bartlett (1970), the “logic of structure”—and henceforth here, the “constitutive structure”—of a specific system describes those relations which are ideal conditions of the possibility of the system, preconditions of that system’s fundamental functional organization—in other words, relations that serve as structural presuppositions. As we shall see later in this chapter, such ideal conditions provide the basis for the identifiability of objects from the standpoint of that system.

If structural presuppositions comprise the relational basis, the “scaffolding,” that enables any functioning system to maintain its structural organization, they are complemented by presuppositions which it makes sense to call specifically ‘systemic’ in nature.

## 2. *Systemic presuppositions*

To define systemic presuppositions we require a working definition of ‘system’. Much of traditional philosophy has been concerned with analyses of the relation between the knower and the known, between truth and a reality thought to correspond to it. Among many Anglo-American philosophers, this has led to a truth-functional approach that centers attention on the relation between individual propositions or statements and that to which they refer. The referential, truth-functional varieties of presupposition discussed in the preceding section reflect this orientation.

The truth-functional, statement-based approach to the logic of presuppositions might be thought to parallel the classical mechanics of particles, the now-outdated “corpuscular” or “atomistic” understanding by physicists of interactions among discrete physical particles, objects, and events. In physics during the last century, this corpuscular approach has been replaced by a fundamentally changed perspective and set of methods centered around the concept of *field*.

Before Clerk Maxwell, people conceived of physical reality—insofar as it is supposed to represent events in nature—as material points, whose changes consist exclusively of motions.... After Maxwell they conceived physical reality as represented by continuous fields, not mechanically explicable.... (Einstein, 1934, p. 65)

The field concept has proved itself to be a conceptual advance over the atomistic approach of classical physics As Max Planck observed:

[I]t is impossible to obtain an adequate version of the laws for which we are looking, unless the physical system is regarded as a whole. According to modern mechanics (field theory), each individual particle of the system, in a certain sense, at any one time, exists simultaneously in every part of the space occupied by the system. (Planck, 1933, p. 24)

The field concept is central to the concept of system I employ in this work. By a field we understand a highly general structural concept relating to ordered wholes, that is, relating to events or ensembles of objects that are recognized to form totalities, whose constituents if they are regarded in the classical sense of discrete, individual “parts” or “components” can very often not be effectively studied or understood, as we shall see. The field concept is an essentially relational understanding of the organization, the ordering, of integrated totalities—which is to say, *systems*. The field concept, especially in the present work, is basic to an understanding of systems. When we recognize the field-based nature of wholes, we abandon a focus upon isolated entities and consider instead the interactional principles of organization that constitute the rules of order upon which organized totalities depend.

The structural presuppositions of a system mentioned in the preceding section are required on a fundamental level by the functional organization of a given system. The phrase ‘functional organization’ is intended in two specific senses: On the one hand, we have the sense which is found in functioning biological systems, functioning computer software and hardware, functioning mechanical systems, etc. In these contexts, the term ‘functioning’ expresses the recognition that a particular system has a *structural* organization that enables it to define or to perform specified functions, to realize specific purposes. On the other hand, the *dynamic* nature of a system, its set of organizing principles that underlie the system’s ability to maintain its organizational integrity as a totality, may be understood functionally in the mathematical sense, that is, in terms of parameters that define its structure.<sup>116</sup> The “organizational integrity” of a system, in general systems terms, involves such properties of systems as hierarchic structure, stability/homeostasis, purposiveness, self-maintenance, etc. The study of such systems-based properties has led to the development of general systems theory itself, and has contributed to the rapid evolution of computer science, information theory, decision theory, game theory, etc.

---

<sup>116</sup> A useful point of entry for philosophers interested in general systems theory remains Laszlo (1975), who follows this course by defining a variety of systems (natural, physical, biological, social, cognitive, etc.) in explicitly functional-parametric terms.

Directly relevant to philosophy is the fact that many philosophical positions meet the requirements of systems as understood here: They are formulations of views and arguments that exhibit organizational integrity, some to a greater and some to a lesser extent. Considered as organized sets of interrelated and interdependent assertions and arguments, philosophical positions often qualify as functional organizations in the first of the above senses of the word ‘functional’, and sometimes, when systematically formalized, also in the second. And, to be sure, there is an extended sense in which philosophical positions frequently function, in their capacity as systems, as extensions and expressions of a philosopher’s individual sense of personal/intellectual identity. Philosophical positions considered in this “person-based” way have been studied in some detail, e.g., by Henry W. Johnstone, Jr.<sup>117</sup>

Systemic and implicitly entailed structural presuppositions together point to a general interrelated class of presupposition that is distinct, one might say, not as a separate species of presupposition, but as a separate *genus*. Such presuppositions do not lend themselves, as we shall see, to effective individual statement-based truth-functional analysis, but require a radical change of approach and of methodology.

To my mind, one of the clearest ways both to understand structural and systemic presuppositions, and to underscore the need for a fundamental change of perspective in connection with these types of presupposition is to consider the work of Leonard Nelson and Humphrey Palmer, both of whom have argued that traditional Kantian and post-Kantian transcendental arguments run headlong into certain peculiar problems of logical circularity. We shall discuss certain of these problems in greater detail in the next chapter when we examine the logic of transcendental argumentation, but here, while we are concerned specifically with the logic of presuppositions, it is appropriate to anticipate some aspects of that discussion.

#### 8.4 The Epistemological Loop

Consider first a set of reflections by German philosopher and mathematician Leonard Nelson (1882–1927). Nelson was critical of some aspects of the logic applied by Kant in his *Critique of Pure Reason*. Since Nelson’s thought is not widely known, I quote from him at length:

Let us look ... closely at the logical character of Kant’s Transcendental Proof. The basic idea is this: these principles [of

---

<sup>117</sup> For example, in Johnstone (1964 and 1970), as well as in others of his publication.

pure intellect] are conditions for experience to be possible, so they apply to every *thing* that could possibly be experienced. Put like this [and in the form of an argument], the syllogism has only one premiss. There must be another one somewhere, for us to draw the stated conclusion with.... An attempt is indeed made [by Kant] to introduce as second premiss an appeal to the reality of experience. We undoubtedly do have real experience. That is a fact, and it seems we have only to state it as a fact in order to infer that the conditions which make this fact possible are actually fulfilled. What is real must be possible. Experience is real; so it must be possible; so the conditions on which its possibility depends must be fulfilled....

If we assert the 'reality of experience' as a fact, all we can mean is that certain empirical judgments of experience have occurred. We do certainly make empirical judgments. But all we can say of them as given fact is that they claim validity. And from the fact that they claim validity we cannot infer that their claim is justified. But *we should have to assume it to be justified, before we could reason back to the fulfillment of conditions necessary to the validity of those judgments.* Now we must give reasons for them before acknowledging their claim to validity, and these reasons would have to consist of the presuppositions on which the validity of those judgments depends. But these presuppositions are just what the Transcendental Proof was supposed to establish. So we should have had to prove these propositions first, before making use of that alleged second premiss about the reality of experience....

If we analyse this process [that of analysing experience] and uncover its presuppositions we arrive at the principles of metaphysics. This, however, does nothing to prove or justify these propositions. All it proves is that we do take them for granted when making empirical judgments. This shows that they claim validity, as any empirical judgment must; it does not show that this claim is justified. So the whole enquiry, properly understood, is purely factual (*quaestio facti*) and not as Kant wrongly suggests epistemological (*quaestio iuris*). *If we try to turn this regressive indication of metaphysical principles into a proof the result is always circular.*

To succeed, a proof of this sort would have to deduce these propositions logically from propositions of some other science; and then they would become theorems of that science. The premisses for proving metaphysical principles would in that case be on loan from some non-metaphysical science, of either the rational or the empirical variety. (Nelson, 1970, pp. 197, 198, 199-200; italics added)

Nelson has done an ample job here of expressing his thought, for which, given our purposes, there is no need for detailed commentary. It is, however, important to underscore the fact that Nelson would like, intends, or presumes—as Kant very likely also wished to do—to cast the Transcendental Proof in quasi-syllogistic form. When Nelson does this, as he relates in the above passage, *circularity of a peculiar sort* comes about. For many readers, precisely why that circularity comes about and what its nature is may not be altogether clear. We'll explore this more fully in what follows.

Contemporary British philosopher Humphrey Palmer has also had a strong interest in the logic of transcendental argumentation, and like Nelson, Palmer has directed attention to the peculiar circularity that is involved when the logic underlying transcendental arguments is exposed to light.<sup>118</sup> His way of examining this circularity tends to be clearer and more systematic than Nelson's.

In a series of publications<sup>119</sup> Palmer has developed two related concepts, that of “presumptive-circularity,” which he abbreviates as “p-circularity,” and that of “backwards arguments.”<sup>120</sup> He has applied these, for example, in analyses of Descartes' *Cogito* and Kant's transcendental epistemology. For our objectives here, I summarize Palmer's approach de-coupled from its historical applications.

For Palmer, a presumptively circular argument comes about when one or more of the premises of an argument cannot be established without drawing the argument inescapably into a loop of circular reasoning. This happens when the conclusion of such a p-circular argument must itself be relied upon in order for one or more of its premises to be established. The justification of such an argument, then, results in unavoidable circularity. In Palmer's view,

---

<sup>118</sup> See Palmer (1981, 1983, 1985, 1994). Palmer was also the translator into English of Nelson's (1970 and 1971).

For readers interested in an earlier recognition of such circularity, see Griffiths (1969, p. 170).

<sup>119</sup> See previous note.

<sup>120</sup> In Griffiths (1969, p. 170) we find very similar observations, anticipating Palmer's, but not developed.



such a p-circular argument is not formally invalid, but he argues that it cannot prove anything; this assessment, he claims, extends to Kant's transcendental deductions and generally affects the majority of transcendental arguments.

In Palmer's view, p-circularity is a variety of *fallacy* that comes about whenever an argument, instead of moving deductively in the normal logically "forward" direction, in accordance with rules of logic, to its conclusion, instead reasons "back" to a presupposition that is taken for granted in arriving at one or more of the argument's premises. (Palmer, 1985, p. 155)

To take an example, suppose that

- (1)  $S$  presupposes  $P$  ('presupposes' here is taken in the traditional truth-functional sense that  $S$ , if it is a referring statement, cannot be true or false unless  $P$  is true or holds),
- (2)  $S$  is a referring statement, therefore
- (3) presupposition  $P$  is true or must hold.

Now premise (1) depends for its truth on the conclusion (3), for  $P$  is a necessary condition in order for  $S$  to be true or false, while (2), which states that  $S$  is a referring statement (i.e., which can be either true or false), also relies on the truth of the conclusion (3). Such an argument involves, Palmer says, a kind of "backward" inference. That inference is not invalid, since for  $S$  to be a truth-functional referring statement,  $P$  must be true or hold; this follows directly from the meaning of 'presupposes' in the above argument. And so Palmer concludes (N.B.: in his terminology, a statement has "propriety" if it is a truth-functional referring statement):

An argument back to a presupposed item from the propriety of the statement presupposing it is bound to be p-circular, as the alleged propriety consists simply in that item being available for reference. Thus one premiss declares that the other can only be established by means of the conclusion those premisses were supposed to be establishing. (Palmer, 1985, p. 50)

Consider another example:

1. The statement "I am (now)" requires permanent external things, as clocks.
2. "I am (now)" is undeniable.

So 3. There are external things.

Here 2 could not be known before 3, if 1 is true, so there is built-in circularity. (p. 81)

The usual “forward” progress of a logically valid argument does not, Palmer claims, take place here; instead, we again confront an example of “reasoning back” from an argument’s conclusion to premises that stand in need of the justification that is provided by the conclusion itself. Palmer tells us:

Given the truth of *P* as a necessary condition for the propriety of *S*, we can rely on the propriety of *S* as sufficient condition for the truth of *P*. But ‘the propriety of *S*’ is not available, except by seeing what *S* presupposes and verifying that presupposition as correct. (p. 90)

As a result, reasoning back in this way, according to Palmer, is p-circular and does not result in proving anything.

A final example:

Argument A is definitely O.K.

Argument A relies on principle P

So principle P must be acceptable.

The latter ... is p-circular, since the principle would need confirming in the course of making sure that argument is O.K. (p. 101)

Palmer therefore concludes:

In an argument from presupposition the second term, *P*, is ingredient in the main term, *S*. This means that *S* cannot be confirmed without first making sure of *P*: after which, argument back from *S* to the truth of *P* is quite superfluous. For this reason all such ‘backward’ inference is bound to be p-circular.... And all transcendental arguments are backward arguments, for they all start from some statement or science or communication-system as given in good working order, and work back from it to the conditions it is alleged to presuppose. (p. 151)

In a later work, Palmer sums up the conclusion to which this reasoning brings him:

The whole project of epistemology thus seems irredeemably circular.

You can't begin, in this critical game, unless you have begun. So you really can't begin at all. To evade this conundrum, people resort to pretense. Keeping ready to hand all the needful apparatus of concepts and standard ways of judging and fundamental certainties, they kid themselves into thinking they have abandoned everything and are beginning intellectual life afresh. But it isn't an entirely fresh start. The slate may be clean, but the hand already knows how to write on it. (Palmer, 1994, p. 37)

Palmer's reasoning is clear, and it succeeds in making explicit and more systematic what Nelson may perhaps have had in view in the passage from his work quoted earlier.

Now, both Nelson and Palmer alike recognize that in the epistemological attempt to justify knowledge, presuppositions must inevitably be relied upon in order to get the whole epistemological enterprise moving. However, the "presuppositions" which they have in mind are not, as I shall make clear, of the usual, familiar truth-functional kind associated with referring statements, but they belong to an altogether different genus, unrecognized by Nelson and Palmer. Reliance upon those "presuppositions"—*if* they were to be of the normal statement-based, truth-functional sort and placed within a syllogistic framework—draws reasoning into the loop that Palmer calls presumptively circular: It is the peculiar kind of circularity which, in Palmer's view, is involved in *attempting to justify that which has to be relied upon in order to make the justification possible*: Reasoning is then pulled into an "inevitable circularity of justification" (Palmer, 1998, p. 84). Nelson, as quoted earlier, describes this curious loop by asserting that, if we are seeking a justification for the claim made by empirical judgments to be valid, we are forced to assume that claim to be justified "before we could reason back to the fulfillment of the conditions necessary to the validity of those judgments." For both Nelson and Palmer, in attempting to provide justifying reasons, we are forced to rely upon, to enlist, the very "presuppositions" upon which the validity of such judgments depends (the very "presuppositions" that Kant wished to establish). A loop inevitably forms as we attempt to resort to a proof that is, in

Nelson's words, "regressive," or in Palmer's words, "reasons back," a proof whose result "is always circular" or is "presumptively circular."

### 8.5 Conceptualizing the Epistemological Loop

If we step back from Nelson's and Palmer's reflections, we note that the model of reasoning they both apply and in which their reasoning takes place is the standard, classical "atomistic" model, expressed in the form of statement-based, premises/conclusion syllogistic logic: It is a model of reasoning that is formulated in terms of truth-functional claims, claims whose truth-functional presuppositions likely would consist, if we were to break down their arguments in detail, of several well-known distinguishable varieties, but all of which belong to a single genus. Referring back to our earlier list of 13 varieties of presupposition, we would suspect that Nelson's and Palmer's model of reasoning accepts—and, most importantly, limits itself to—the meaningfulness and legitimacy of the first 11 varieties of presupposition, that is:

1. presuppositions as preliminary assumptions
2. conditional presuppositions
3. presuppositions as logical premises
4. presuppositions as logically antecedent suppositions
5. presuppositions as the ingredients of definitions
6. presuppositions as suppositions of language use
7. presuppositions of pragmatical activity
8. presuppositions of missing premises
9. presuppositions of existence
10. linguistic presuppositions of reference
11. presuppositions of concepts

Not all of the above varieties of presupposition play explicit roles in Nelson's and Palmer's analyses of the epistemological loop (numbers 3, 5, 10, and 11 evidently shoulder much of the burden). But it is not an unlikely conjecture that were we to imagine engaging in person-to-person conversation with Nelson and Palmer about their views concerning the epistemological loop, it would feel quite natural for both thinkers to rely upon the familiar traditional vocabulary consisting of the varieties of presupposition drawn from—but limited to—the above list.

The accustomed syllogistic, truth-functional, statement-based model of reasoning that Nelson and Palmer apply in seeking to understand what I've called the 'epistemological loop' misses the point: That model of reasoning, much like the discarded atomistic approach of classical physics, is inadequate

if we wish to conceptualize the *internal limitations* of the epistemological situation we face. *A different model of reasoning is required to understand why and how the epistemological loop comes about to begin with.*

It is important to note that both Nelson and Palmer point to what they judge to be a need for justification originating “outside” the loop—that is, if epistemology is to produce a successful proof that justifies its own possibility and validity. We recall that Nelson concluded

To succeed, a proof of this sort would have to deduce these propositions logically from propositions of some other science; and then they would become theorems of that science. The premisses for proving metaphysical principles would in that case be on loan from some non-metaphysical science. (Nelson, 1970, p. 200)<sup>121</sup>

Palmer similarly remarked:

Can we establish, say, the Law of non-contradiction by showing that even those who would deny it must perforce rely on it? We can show that they should not deny it; and maybe no one else is going to. But does this prove that it is true? Such a proof would require a starting-point *independent* of the item being proved. (Palmer, 1994, p. 39)

In Palmer’s view, in order to break out of what I’ve called the ‘epistemological loop’, formed by an argument consisting of a set of premises followed by a conclusion that must be appealed to in order to justify one or more of the premises, it would be necessary for the argument’s proponent somehow to “show that the premisses are really available to him when the conclusion is not” (Palmer, 1985, p. 156). By ‘available’ Palmer means that we should somehow be able, *independently* of the information contained in the argument’s conclusion, to verify the truth of the premises. This of course does not happen when there is an inescapable “backward” inference embedded in an argument like the following: If *S* presupposes *P*, and if *S* is a referring statement, then, the argument concludes, presupposition *P* must be the case or

---

<sup>121</sup> In developing his own post-Kantian position, Nelson eventually came to embrace what he called “immediate knowledge,” which he believed avoided both the circularity problem and the need for independent supporting justification. Cf. Nelson (1971, pp. 171-187). It would take us too far afield to consider the position he defended, one that rests on an intuitive “feeling for truth” (Nelson, 1970, p. 7) that does not satisfy our more rigorous standards here.

hold. Such an argument, Palmer claims, is “incompetent as proof” (Palmer, 1985, pp. 41, 90). What is required, he reasons, is that there be some *independent* way of recognizing (knowing or verifying) that *S* really is a referring statement without relying upon the truth of the argument’s conclusion. But, given the epistemological loop we are pulled into when we seek for a way to self-justify knowledge, there is, in Palmer’s view, no way to do this without fallacious circularity.

What Nelson and Palmer have come to bump up against are *the limits of epistemological self-justification*. Their claim that there is a need for justification that originates “outside” the loop, or “independently” of the loop, is really, as we shall see later on, a *symptom* of having hit a *metalogical barrier*, one that, in principle, we cannot go beyond without importing the very ground that we would seek in this way to justify. As we shall see later, it is due to this metalogical fact that Nelson’s and Palmer’s belief that transcendental argumentation requires “independent” evidence or justification in order to break out of its peculiar circularity *cannot*, in principle, be met.

Here lies the importance of Nelson’s and Palmer’s efforts: They indirectly reveal to us *epistemological boundaries* that cannot, in principle, be exceeded. The study of such boundaries forms the core interest of this study.

Where, then, does this leave us? The answer to this question, as I see the matter, lies in the need to recognize the way in which structural and systemic presuppositions function.

When we come to see that the task of accounting for the possibility of knowledge entails and can only take place within a *dynamic*, interrelated general *system*, the presuppositions that we need to attend to are then seen no longer to belong to the general statement-based, truth-functional referring genus, but are presuppositions of an essentially different kind. They are not capable of being demonstrated directly, by applying statement-based premises in a syllogistic argument that leads to a conclusion. They can only be demonstrated by what I’ll call ‘*destructive testing*’.

How do we know that a particular cable of a suspension bridge is *necessary* to support it? We may sit at a desk and computer and calculate forces, but the acid test is to cut the cable to see if the bridge then collapses. How do we know that a pumping heart is necessary to sustain human life? We see the result that follows complete heart failure. How do we know that a line of computer code is necessary for the functioning of the larger program in which it forms a part? We delete it and see whether the program is still capable of running effectively. In all of these cases, different kinds of systems are involved. The conditions that must necessarily be met in order for any of these systems to function are structural/systemic conditions that comprise presup-

positions that are not of the classical, statement-based, truth-functional kind.

The conceptual system that is in view when we ask the epistemologist's central question, How is objective knowledge possible?, is no less a general dynamic system, an integrated interrelated whole whose elements mutually depend on the others and cannot be what they are except in their relations to those others.

When we ourselves form part of that general system, how is it possible, if it is possible at all, for us to self-justify the system in which all of our tools of justification form integral parts? The answer lies in destructive testing: We test a purported precondition by denying it in order to discover whether the general system—in a specific sense that we will examine in detail later on—undergoes collapse.

Palmer, I believe, began to touch the edge of this recognition when he expressed thoughts like the following:

The statement that A presupposes B is always *relative to some system* of thought or of ideas. Ingredience is a feature of the *system*, not just of the ingredient concept and that which presupposes it. We commonly neglect to mention the *system*, because it is so obvious. (Palmer, 1985, p. 98; italics added)

A principle is fundamental to some *whole system* if discarding it means destroying that *system as a whole*. The fundamental character of a principle is thus decided *by reference to the thought-system*, the 'science' based on it, and not by reference to the scientists using it, or to the age in which they live. (p. 171; italics added)

Yes, because it is so obvious, we do commonly neglect not only to mention the system, but we neglect to study the logical consequences that follow from the fact that we are dealing with a system—but not only that, we fail to realize that when dealing with a genuine integrated dynamic totality, the logical tools that are applicable are not to be found in statement-based, syllogistic, truth-functional argument, or in the associated categories of presupposition whose logic has traditionally been studied.

In later chapters, I develop a group of logical techniques that are appropriate when dealing with such systems, tools that are designed to reflect our recognition of the internal limitations that come inescapably with systems in which we, our thought processes, and capacities of expression form integrated

and interdependent constituents. In this context, the concept of field will be central.

### 8.6 Presuppositions of identification, continued

Presuppositions of identification are properly a sub-class or species of structural/systemic presuppositions. They comprise preconditions of the identifiability of objects of reference relative to their associated frameworks of reference. They are presuppositions that are specifically framework-relative and are the foundation of the metalogic of reference developed here. Presuppositions of identification describe the conditions without which the identity of an object of reference cannot be recognized. They are theoretically the most fundamental, general, and ubiquitous variety of presupposition in no matter what area of human thought or discourse we may wish to consider.

We routinely recognize a number of forms and degrees of identification, among them:

- *Unique specification*: e.g., one's Social Security number, a researcher's ORCID identifier, a telephone number, a physical address, the exact GPS coordinates of a location—all serving the purposes of uniquely specifying a particular person, a single telephone line, or a specific geographical location.
- *Identification according to rule*: e.g., the 47th prime number; the result of multiplying  $\pi$  by itself a certain number of times; a set of instructions to find a treasure, attached to a map to which those instructions apply.
- *General identification*: e.g., designation of an object's class membership, type, species, etc., serving the objectives of less individually specific identification by indicating inclusion in a group.
- *Abstract identification*: e.g., the algebraic definition of 'function', the definition of 'valid proof' in a system of mathematical logic, the concept of number or of hypothetico-deductive theory—examples of formally defined operations, rule-based idealizations, or methodologies.
- *Vague identification*: e.g., specification through approximations, ambiguity, or resemblances, which can make it possible to narrow down,



but only within a fuzzily defined range of reference, identifications that are only indefinitely determined.

No matter what we attend to, seek to communicate, recognize, or intend to designate, define, describe, or allude to, there we find forms and degrees of identification, some very specific, others general, some abstract, some vague or ambiguous. By ‘identification’ and the phrase ‘identifying reference’ I therefore have in view not only the most commonly assumed meaning that involves unique specification of an individual object, but rather the very broad and abstract sense in which identification/identifying reference obtains when that which is described is determined as that to which there is reference. Such broadly understood identification may range from unique to general, from vague to abstract, or consist only in specification (precise or vague) according to rule.

We normally think of identification as a process, an activity that involves the application of criteria, a set of standards, a procedure that produces a result in the form of an “identification”—much as a bird watcher, consulting his field guide, applies a set of defining specifications to an observed bird and concludes that it is of such-and-such a sort: his observations result in “identification.” Identification in this prevalent sense involves perceptual and cognitive selection or discrimination skills, abilities to search and match criteria, make connections between perceptual data and reference samples, apply classification categories, and detect commonalities and differences.

We recall that this common view takes on a similar form with respect to the concept of reference, as discussed in {3.4}. There we noted that the “psychology of referential thinking” is laden with philosophically problematic presumptions that characteristically include unreflective and unanalyzed notions of the self, its activity, its relation to other things and other minds, the belief that a study of the grammar of ordinary language is an important source of epistemological understanding, etc. And so it is also in connection with the “psychology of identification”: The pragmatic, activity-based conception of reference parallels the activity-based view of identification as a psychological activity that results in the human cognitive recognition of the identity of objects.

This is the familiar, epistemologically naive view of identification, but it is not what here is meant by the term. To make appreciable progress in epistemological understanding, we need to look at the subject in a radically different way, in a theoretically fundamental and more systems-responsive manner, in which the “object identified” and the “object’s identity” are inextricably fused.

As we shall see in detail later on, objects have the identity they have as a function of the identification which is possible by virtue of the identification frameworks that are necessary in order to identify them. This, considered from the standpoint of systems theory, is the source of the epistemological loop that misled Nelson and Palmer. Once we realize that identity and identification are two aspects of the same framework-relative logic, we have begun to enter the area of study that I term the ‘metalogic of reference’.

Identification and the identity of the object identified are distinguished only by virtue, in the case of identification, by our stressing the primary importance of the role of the identification framework, and in the case of the object’s identity, by our stressing the primary importance of the object’s identity, often to the neglect of the framework in terms of which the object has, and can have, the identity that is identified.

From this standpoint, presuppositions of identification are understood in the following way: If *S* is an identifying description of an object *o*, and *P* is a condition that must be met in order for *S* to obtain or succeed in identifying *o*, then ‘*S* presupposes *P*’ means that for *S* to identify *o*, *P* is a necessary condition of *o*’s identifiability. Such a *P* is what is meant by a presupposition of identification.

We may distinguish two varieties of such conditions of identifiability:

- object-related: having to do with identification criteria that objects of reference must satisfy in order to be identifiable in principle, and
- system-related: having to do with the frame(s) of reference in terms of which identification can obtain

Conditions of identification are not themselves true or false: they are satisfied or they are not. They form the basis of identification; when this basis is not provided, identification is not possible.

Conditions of identification, and hence presuppositions of identification, are structural/systemic rather than truth-functional in nature. They render *possible* that which, without them, would collapse—much like the cables that support and prevent a suspension bridge from falling.

We could of course choose to specify that such conditions must hold, by means of asserting the truth of statements that say this, but this mode of expression is at one remove from the subject with which we’re concerned. It is clearest to remain on a level that informs us and reminds us explicitly that we are dealing with structural/systemic conditions of identification, rather than properties of statements and relations between statements.

Even with this clarification, since our study of presuppositions of identification is expressed in a natural language, and the results we reach will be expressed in that language by means of assertions that claim to be true, it is easy to slip into a mistaken view that we are therefore dealing with linguistic matters. We are not: We are dealing with underlying logical structures that are non-linguistic. The difference between these two levels of theory may be made clear in the following way:

When necessary conditions of identification are not met, identification cannot occur. In this sense, from the standpoint “closest” to the subject-matter in view, the underlying metalogic of identification is *bivalent*: either these conditions are satisfied, or they are not.

However, when language enters in and assertions in language are made that deny the conditions of identification that must be presupposed for a position’s or a speaker’s intended reference to be possible, then statements of such a kind are neither true nor false, but meaningless. To make a claim that essentially amounts to this—“Reference is made to that to which reference is impossible”—is to say something that is self-undermining and nonsense. The logic required by this language-based expression is therefore *trivalent*, not bivalent. We shall develop this in some detail later, in {11} and the Supplement.

The metalogic of presuppositions of identification as described here is essentially modal: It is appropriate to call such presuppositions ‘metalogical’ in the sense that they comprise conditions that must logically be satisfied before an object *can possibly* be said or thought to have an identity. They are, in this sense, conditions, often called ‘*preconditions*’, that pertain to formal ontology. By “formal ontology” I mean the theory of formal properties and formal relations that are fundamental and shared in common by a wide range and variety of identifiable objects—properties and relations that are invariant, or may be transformed without loss from one reference frame to another, and irrespective of the nature of those objects. Such preconditions that pertain to formal ontology are not, though they underlie, conditions of linguistic discourse. They are the metalogical, formal substructure and systemic framework that make identification (and hence discourse involving identification) possible.

Hence, metalogical conditions of identification are neither about statements that are true or false, nor about the contextual beliefs held by speakers and hearers. They are, instead, formal, structural/systemic principles that constitute the basis of all identifiability.

If person *x* says, “The tree in front of the house has begun to put out new leaves,” the statement that *x* has made presupposes several combined varieties

of presupposition: They include, referring to our earlier list, presuppositions as suppositions of language use (6.), presuppositions of pragmatical activity (7.), presuppositions of existence (9.), and linguistic presuppositions of reference (10.). The statement made by  $x$  presupposes that there is such a tree in front of the house (6., 9., 10.);  $x$  contextually presupposes that he or she believes there is a tree there (6., 7., 10.); but the foregoing presuppositions rest on underlying metalogical conditions of identification, which have to do with the *identifiability* of trees, leaves, houses, spatial orientation in relation to, e.g., the house in question, newness of a tree's leaves, etc. These conditions relate to the objects identified and to the framework of reference in terms of which such things *can* be identified (11., 12., 13.).

If  $P$  is a metalogical condition of the *identifiability* of  $S$ , and if identifying reference is made to  $S$ , then  $P$  must *necessarily* hold or be satisfied; if condition  $P$  is not satisfied, identifying reference to  $S$  is *impossible*. The reader should note the central role in the preceding statement of the modal concepts of possibility and necessity.

Recalling the maximally general concept of possibility described toward the end of the last chapter, presuppositions of identification render explicit those conditions that must be met in order for identifying reference to be possible to those objects to which reference, by virtue of the satisfaction of those conditions, is thereby enabled. These conditions, to speak somewhat metaphorically, form the interstices within which a network of possibilities is determined: They form the network of positive and negative constraints discussed in the preceding chapter that define possible objects of reference relative to a framework. In this sense, they constitute the maximally general, maximally basic substructure of all possible objects of reference.

## APPENDIX TO CHAPTER 8

### Rule-based Games and Passmore's and Collingwood's Presuppositions

The technique, before the advent of calculators, that was once taught in some schools which makes it possible manually to extract square roots is a technique that presupposes basic operations of arithmetic (division, multiplication, addition, subtraction). This manual step-by-step procedure cannot be formulated without taking these basic arithmetical operations for granted. The manual extraction of square roots is an example, in miniature, of a procedure that presupposes a certain set of rules that can be considered to comprise an abstract game.

The rules that define a game may be simple or complex, and when the rules are numerous and interrelated, the game they define can become comprehensive enough so that the game is no longer a delimited set of permitted and prohibited moves, a game isolated from a broader scope of human concern, but rather it can determine a general approach, an orientation, or a method, which may take the form, for example, of a scientific theory, a methodology, or a profession's or a society's "domain of discourse." All can be thought to represent games in this general sense.

The transition from a delimited game to the broader, more encompassing domain of a theory or area of discourse is not marked by a clear line of separation. This is reflected in the fact that some kinds of presuppositions may foreshadow or even involve other varieties, as can be the case, as we will see here, in connection with presuppositions of concepts that may merge with the varieties of presupposition that I've called 'structural' and 'systemic'. Not all the varieties of presupposition are compartmentalized by strict dividing lines of demarcation; some may overlap one another.

The inventory of varieties of presupposition and a study of their underlying logic in {8} would, by many philosophers, be considered incomplete without some explicit mention of John Passmore's "presuppositions of all proposing" (Passmore, 1961, p. 68) and R. G. Collingwood's "absolute presuppositions" (Collingwood, 1940). These kinds of presupposition that interested Passmore and Collingwood do not lend themselves easily to categorization despite the choice available to us among the 13 varieties distinguished in {8.1}. The presuppositions that concerned both men seem rather to form an

ambiguous transitional or blended variety of presupposition that we might plausibly locate somewhere between presuppositions of concepts and structural/systemic presuppositions. The fact that there is no clear fit for Passmore's and Collingwood's presuppositions is admittedly due to their shared shortcoming in defining clearly what they had in mind. The framework supplied by rule-based games may help to provide some of the clarification which they did not.

In {6.5}, I described Passmore's presuppositions of all proposing, which, as we saw, he equated with what he called 'the invariant conditions of discourse'. We found that his account falls short because he failed to define explicitly what such presuppositions really are, how they can be identified, or how they can be demonstrated. My sympathetic conjecture was that Passmore was moving toward a recognition of general and invariant principles to which a philosophical argument can appeal when it makes recourse to the technique of self-refutation. I placed Passmore's concept of "presuppositions of all proposing" in a "transition area" between pragmatical and metalogical self-referential argumentation.

If we transpose his efforts into a rule-based game-theoretic context, Passmore's presuppositions acquire what I think may be both a little clearer as well as an appropriate application. Transposed to the context of a specific game (for Passmore, an individual philosopher's asserted position), his presuppositions are the rules that a player of that game must accept or take for granted in order for that game to be played; they are rules according to which that philosopher's position is argued, rules that are indispensable to the formulation and communication of that position. For Passmore, such presuppositions in a sense resemble presuppositions of concepts, as when he argued that we cannot renounce the fundamental claim that some of our statements have meaning, or the fundamental claim that at least some assertions must be made in order to engage in discourse. An argument of this kind resembles an example of presuppositions of concepts given earlier in {8.1.11} in which the concept of lying presupposes the concept of telling the truth. In Passmore's argument, the concept of meaning and the concept of asserted truth entail presuppositions which Passmore maintained were undeniable. Such presuppositions function as rules that he believed are indispensable to human discourse.

Viewed from another perspective, such rules define a total system, as in the case of a philosopher's systematically formulated position. Without them, that system collapses. Such presupposed rules have a structural/systemic character in the sense that a philosopher's position can comprise a dynamic interrelated system of propositions with its own logical scaffolding.

Passmore did not provide an adequate account of his presuppositions of

all proposing to establish whether he thought they were wholly or partially presuppositions of concepts, or wholly or partially structural/systemic presuppositions. Given the ambiguity, they appear to have aspects of both.

In a somewhat similar way, R. G. Collingwood's "absolute presuppositions" also have this transitional nature, and they equally lend themselves to a game-theoretic translation. Collingwood claimed that such presuppositions are not themselves propositions; they are not true or false, but must be taken for granted. He claimed that they cannot be demonstrated—"to talk of justifying them is to talk nonsense" (Collingwood, 1940, p. 46). Like Passmore's presuppositions, Collingwood's absolute presuppositions function as rules, which are not themselves "true" or "false," but which define a total game. One of the few commentators to note in passing their resemblance to the rules of a game was Heikki Saari, who suggested that Collingwood's "[a]bsolute presuppositions resemble the rules of a game in the sense that when we participate in the social practices in which they are embedded, we cannot choose whether to make them or not" (Saari, 1991, p. 66).

Unfortunately, Collingwood, like Passmore, did not develop his concept of absolute presuppositions with any logical rigor. In fact, he seemed to balk at the task of defining explicitly what they are or of justifying them. He wrote: "[w]hen I speak of finding out what they are I do not mean finding out what it is to be an absolute presupposition, which is work for a logician..." (Collingwood, 1940, p. 54); "...absolute presuppositions do not need justification" (p. 44); "...to talk of justifying them is to talk nonsense" (p. 46). His notion of absolute presuppositions reduces fundamentally, I think, to the basic unquestioned assumptions made by the conceptual framework that is in vogue or in force at a particular historical time. This is now considered to be the province of anthropology, not of a rigorous philosophical study, for anthropology includes the study of the most basic beliefs which in a particular society are considered to be beyond question.

From the standpoint of game theory, Collingwood's absolute presuppositions, like Passmore's presuppositions of all proposing, function as rules that are indispensable to a game, but whether Collingwood's are truly "absolute" in any sense, or whether Passmore's presuppositions are truly "invariant" in any sense, neither philosopher explained adequately or justified.

## 9

# Transcendental Argumentation and the Metalogic of Reference

[I]t has the peculiar character that it makes possible the very experience which is its own ground of proof, and that in this experience it must always itself be presupposed.

– Immanuel Kant (1965/1929, B 765)

Already in this book I have used the phrase ‘transcendental argumentation’ on numerous occasions, often in passing—in connection with the work of Kant, Husserl, Reichenbach, Isaye, Strawson, etc., and, in {6.8}, specifically in connection with metalogical self-referential argumentation. In this chapter, I look more closely at the nature of transcendental argumentation as it has been approached in the past, and then apply to the discussion certain of the results of earlier chapters. In {7} we were led to highly general concepts of possibility and necessity, and in {8.1.12–8.1.13} developed the concepts of structural/systemic presuppositions and presuppositions of identification. We’ll now make use of these concepts to develop further the innovative approach to transcendental argumentation that is the object of this study.

### 9.1 What is transcendental argumentation?

The chief purpose of transcendental arguments is to identify, and then ideally to demonstrate, on a radically fundamental level of undeniability, preconditions of possibility of whatever subject-matter is in view. The motivation of philosophers to accomplish this is a direct expression of one of the most basic reasons that attract people to philosophical study: to reach a level of certainty, theoretic comprehensiveness and generality, and permanence of a non-contingent kind that are unobtainable in the great majority of disciplines, mathematics standing as the most familiar exception.

In this chapter’s opening quotation from Kant’s *Critique of Pure Reason*,



Kant was speaking of the principle of causality. The very language he employed points to the “peculiar” variety of circularity which, as we saw in the last chapter, concerned Nelson and Palmer, a kind of circularity that seems to come about when one looks for ways to justify the very basis of epistemological justification.

Depending upon how it is defined, transcendental argumentation will often exhibit this form of circularity—or ‘recurvature’, as I prefer to call it with a bow to topology, in order to distinguish it from the variety of vicious circularity familiar in traditional logic. Much that is central to understanding such recurvature has, in my estimation, largely escaped philosophers who have engaged in transcendental argumentation, and yet glimmerings of its distinctive and unusual logic appear now and then in their thought.

We’ll begin by looking briefly at several of the main ways in which transcendental argumentation has been approached. I do not propose to summarize the history of transcendental argumentation or comment on the work of individual philosophers, which requires a book in itself, of which several have been written.<sup>122</sup>

In general, transcendental argumentation involves a defining context, a focus of interest and analysis which seeks in one way or another to identify preconditions of possible experience, preconditions of possible truth, preconditions of possible knowledge, preconditions of possible justification, or preconditions of possible meaning. The role of the repeated words ‘preconditions’ and ‘possible’ is of course fundamental, and much depends on how those central terms are understood. Again in general terms, such transcendental “preconditions” relate to presupposed principles or conditions that are judged *necessary* in order for experience, truth, knowledge, justification, or meaning to be *possible*. Few philosophers who have engaged in transcendental argumentation have explicitly developed a general theory of the presupposed concepts of possibility and necessity, while what philosophers mean by ‘preconditions’ varies according to their focus of interest. And it is probably fair to say that even fewer philosophers have devised ways to *prove* that the alleged preconditions they have pointed to really are necessary preconditions without which a subject-matter becomes impossible. We cannot expect a high degree of self-conscious analytical clarity in any area of study that easily encourages highly abstract and sweeping generalities, and the theory of transcendental argumentation is no exception.

Three main approaches to transcendental argumentation may be distinguished here: The first especially bears the imprint of Kant. It is a theoretic-

---

<sup>122</sup> See, for example, Bieri, Horstman, and Krüger (1979), Stern (1999, 2000), Stapleford (2008).

cally highly general approach that can be characterized by its very broad interest in determining “the principles without which no object can be thought at all...” Such principles are, for Kant, *necessary* in the sense that “no cognition can contradict [them] without at the same time losing all content, i.e. all relation to any object, hence all truth.”<sup>123</sup> Transcendental arguments with an objective like this seek to derive from premises concerning human experience, thought, or knowledge a conclusion that identifies the necessary conditions of those premises—much as we saw in the previous chapter’s discussion of Nelson’s and Palmer’s analyses of transcendental argumentation. Kant developed arguments, incorporated in his transcendental deduction, that sought to demonstrate such necessary conditions of experience, thought, and knowledge. This approach to transcendental argumentation is explicitly human-centered, concerned with human experience, human cognition, human knowledge, and with the principles without which our experience, thought, and knowledge would not be possible or in some sense coherent.

A second approach to transcendental argumentation, which has come to dominate the interest of many philosophers during the past several decades, has concentrated on arguments against skepticism, seeking to show that the skeptic *cannot possibly* in a coherent or meaningful way articulate his or her position without at once granting conditions that the skeptic wishes to deny. Some of these anti-skeptical arguments take the evident form of pragmatical self-referential arguments (see {6.2}).<sup>124</sup> Similar transcendental arguments have been proposed against those who deny that other minds or independently existing material objects can be known.<sup>125</sup> In all of these contexts, the purpose of such arguments is to point to the necessity of granting certain preconditions of truth, language use, conceptual meaning, or justification which a criticized position denies or considers to be doubtful.

A third approach to transcendental argumentation has been developed by the author: It employs a distinctive and radically different process of transcendental argumentation to show that there is a need for critical revision of much thought, reasoning, and beliefs which are accepted and propounded—whether by common sense, by scientists and other professional groups, or by many philosophers. To do this, in a series of publications I developed a systematic and systems-sensitive method that makes it possible to identify and to prove that metalogically self-referentially inconsistent thought, reasoning, and

---

<sup>123</sup> In these passages Kant was referring to his transcendental analytic (Kant, 1900/1887, A62-3/B87).

<sup>124</sup> See, for example, Rehg (1989) and Bardon (2005).

<sup>125</sup> See, for example, the useful “Bibliography of Works on Transcendental Arguments” by Isabel Cabrera, covering the period 1939-1998, may be found in Stern (1999, pp. 307-321).

beliefs both are pervasive and are rationally unacceptable because they are metalogically self-undermining—that is, they implicitly or explicitly reject preconditions that must be granted in order for such thought, reasoning, and beliefs to possess possible meaning.<sup>126</sup> Such an approach has two main objectives: negative, in its capacity as a tool of criticism, and positive, both in its revisionist objective to provide meaningful replacements for unacceptable views, and in its ability to render explicit the conceptual boundaries of possible meaningfulness. This third approach forms the subject-matter of the present study.

Of these three approaches to transcendental argumentation, the first two have suffered from three major shortcomings:

(i) their general inability, inadequacy, conceptual ponderousness, or obscurity in proving that alleged “transcendental preconditions” really *must*—in some presumably strong meaning of *necessity*—be granted in order for a set of statements, propositions, concepts, or a position or theory that is/are under analysis *possibly* to function as intended, to be true, valid, to refer, to be meaningful, etc.;

(ii) their perceived circularity of argument, which renders them questionable or ineffective, a peculiar circularity that results from a Procrustean forcing of transcendental argumentation into a traditional, syllogistic, statement-based logic; and finally, as we shall see later,

(iii) their acceptance and reflectively unselfconscious employment of concepts which themselves are metalogically self-undermining. (In order to give the reader an early example here: Transcendental argumentation has commonly been used in an attempt to gain what is believed to be epistemologically or metaphysically justifiable access to so-called ‘transcendent’ objects; we shall see that, from the standpoint of the metalogic of reference, such “access” is impossible and meaningless.)

All three approaches to transcendental argumentation involve a general form of reasoning that proceeds in the following steps that should by now be familiar: A subject-matter is specified; it may be a statement, a proposition, a

---

<sup>126</sup> See works by the author listed in the References.

philosophical position, a scientific theory, or a set of beliefs. That subject-matter is reflectively analyzed with the goal of identifying necessary conditions without which that subject-matter would, in some important and basic sense, collapse—rendering it impossible for the initial statement, proposition, position, theory, or beliefs to be true, valid, self-consistent, meaningful, or justifiable, etc. Syllogistically expressed:

(1) If  $A$  is to be possibly true (valid, self-consistent, known, meaningful, justifiable, etc.), then precondition  $B$  must necessarily hold:

$\Diamond A \rightarrow \Box B$

(2)  $A$  is the case:  $A$

(3) If  $A$  is the case,  $A$  must be possible:  $A \rightarrow \Diamond A$

(4) Therefore,  $B$  must hold:  $\Box B$

The argument claims that if the necessity of  $B$  is denied (i.e., denying that  $B$  is a transcendental precondition of  $A$ ), then  $A$  is rendered *impossibly* true (impossibly valid, consistent, known, meaningful, justifiable). Beyond this, as Nelson and Palmer would wish to point out,  $B$ —if it truly functions as the indicated precondition—already must hold in the very recognition of (1).

## 9.2 Transcendental argumentation as structural/systemic

In the previous as well as in this chapter, I've suggested that it is a mistake to think that transcendental argumentation consists of genuine "arguments" in the sense commonly understood of a group of premises that lead to a conclusion that is validly derived from them. I have claimed that the syllogistic approach fails to do justice to the structural/systemic nature of transcendental argumentation. What, then, would 'transcendental argumentation' mean if it does not result in statement-based arguments consisting of a group of truth-functional premises followed by a conclusion whose truth is believed to follow from those premises according to the rules of logic to which we assent?

Unless referring to statement-based, syllogistically structured transcendental arguments formulated by other philosophers, I will shift from the use of the phrase 'transcendental argument', replacing it with the phrase 'transcendental argumentation'. The reason behind this choice of words is precisely to avoid the statement-based mould of premises-leading-to-a-conclusion often associated with the term 'argument'. 'Transcendental argumentation', as I will use this phrase, is instead intended to express the dynamic character of an integrated, interrelated system of conceptualization or of thought, a system in

terms of which statements made during the course of transcendental argumentation are assured of the *possibility* of conveying the meaning they serve to communicate. And so the question immediately arises: How, if transcendental argumentation is to be understood in dynamic systems terms, is that understanding to be formulated? Later in this work we shall have an opportunity to see in detail how this is possible; here, I want to give the reader a brief preview:

Instead of thinking that transcendental argumentation consists of “arguments” in the sense commonly expressed (again: premises from which a conclusion is inferred), transcendental argumentation appropriately and logically needs to be understood in terms of inherently *tautological*, *self-validating* descriptions of the logical *dynamic* of a reference *system*. (On self-validation, see {4.6}.) Understood in systems terms, such a description is reflexively self-validating: We cannot isolate a group of “premises” that do not already have built into them the “conclusion” we seek to establish: The necessary principles, without which reference in that system is not possible, are *built in from the start*. Furthermore, such a description is tautological in the sense that it does not add information that we did not already possess. Much like the tautology  $p \vdash p$ , if  $D$  is a description of the transcendental preconditions that provide for the possibility of a subject-matter  $S$ , the proposition that  $S$  strictly entails  $D$  is a self-validating tautology: Any attempt to affirm  $S$  while denying  $D$  is metalogically self-referentially inconsistent, while to affirm  $S$  while concomitantly affirming  $D$  is self-validating and tautologous; no information we did not already have is acquired as a result.

If initially this reasoning appears obscure, it is not. It makes no sense when dealing with a referential system, with an interrelated systemic totality that provides a basis for identifying reference, to attempt to isolate a set of statements that have ingredient in them—built into them—the very preconditions that render reference within that system possible, and then to attempt to infer from that set of statements a conclusion concerning those preconditions, a conclusion whose truth must already be structurally-systemically presupposed by that very set of statements. There is no “argument” here in the traditional sense; there is rather a self-validating affirmation (or a metalogically self-undermining rejection) of the system of interrelation that makes reference in terms of that system possible.

That there is an obvious self-referential aspect to such argumentation should be evident to the reader. In fact, as we shall see, its metalogic is—to use an appropriate word again drawn from topology—“recurved,” much like a Möbius strip or Klein surface. It is the logic of systems that have no apparent borders and yet have inescapable limits.

By ‘transcendental argumentation’ I therefore mean argumentation that has two central objectives: (i) to identify the preconditions of reference of a particular system, or reference frame, and (ii) to show that these indeed are necessary preconditions of that system by demonstrating that affirming them is self-validating, and that rejecting them is metalogically self-referentially inconsistent.

### **9.3 Transcendental argumentation: Possibility, necessity, and identifiability**

The reader will recall that the most theoretically general concept of possibility reached in {7} understands possibility as a function of parameters of constraint. Possibility from this theoretically abstract perspective is always a function of some set of constraints; the more such constraints are restrictive, the more limited will be the range of what is possible in terms of the reference system they define; as such constraints are loosened, the conditions of framework-relative admissibility are broadened.

We distinguished “positive” and “negative” constraints, the first functioning, relative to a given framework, as prescriptive rules, which must be followed, and the second as injunctive rules, which express prohibitions. In these terms, we understood that what is *possible*, relative to a given framework, is both what is not prohibited by negative constraints, and what does not conflict with positive constraints. What is *necessary* is what accords with a framework’s positive constraints while complying with the injunctions imposed by its negative constraints. In these senses, possibility and necessity are rule-based and rule-governed framework-relative concepts. They lay down the fundamental requirements (necessities) that must be accepted in a given frame of reference, while they establish boundaries of admissibility (possibilities), again relative to that reference frame.

When, using the variety of transcendental argumentation that is in view here, we identify the metalogical preconditions of reference—that is, the specifically transcendental preconditions—satisfied by a given frame of reference, we not only make explicit what is possible from that standpoint, but what is necessary. Such preconditions of reference, as we have seen in earlier chapters, function to provide the basis for *identifiability*, whether such identification is precise or vague. To be an object of reference relative to a given reference frame *is* to be an identifiable object, one that has some degree of identity: To be a possible object of reference *is* to be an entity with some degree of identifiable identity.

We are not accustomed to acknowledging the essential, the necessary,

connection between an object's possession of identity and the framework(s) that make possible its identifiability. Instead, we often—indeed most often—neglect or deny the relativity of the identity of objects to the reference frames that provide for their identifiability. We shall later have occasion to see, in a variety of contexts, how this neglect or denial results in metalogical self-referential inconsistency.

The modal concepts of possibility and necessity to which we have been led are functionally defined in terms of framework relativity, so that metalogical possibility is understood in terms of framework-relative possibility, which in turn recognizes that metalogical possibility provides the basis for the identifiability of a range of objects as a function of a particular frame of reference.

In these terms, the central objective of successful transcendental argumentation results in a reflective recognition and demonstration that the preconditions of reference which such argumentation makes explicit provide for the identifiability of a range of objects from the standpoint of the reference frame under analysis. That recognition does not come about as a derived consequence of a set of truth-functional premises, but is instead the reflective recognition of a tautologous relation of inescapable relativity of the identity of objects with respect to the frame(s) of reference in terms of which their identifiability is assured. The demonstration that such a relation obtains is realized when either we attempt to deny those preconditions while at the same time presuming that objects to which we wish to refer possess the identifiability which those preconditions make possible, a presumption that results in metalogical self-referential inconsistency, or we confirm, through self-validation, the tautologous connection between those preconditions and the identities of the objects in question.

Our analysis will remain on this abstract and general level until, in Part III of this study, we apply the results of the metalogic of reference to specific examples.

#### **9.4 Transcendental argumentation and meaning**

All three of the approaches to transcendental argumentation distinguished here share some concern, although often only implicit, for the role of meaning in a transcendental study of experience, truth, knowledge, justification, etc. The concept of meaning is relatively seldom the explicit focus of an approach to transcendental argumentation, and then that concept tends to lie implicitly and unexamined in the background. The explicit interest of such argumentation may be the preconditions of the perception of objects, of experience gen-

erally, of the awareness of others, of cognitive thought, of asserting or believing, etc., but inevitably, since transcendental argumentation is committed to formulating its expression using meaningful concepts, employing meaningful reasoning expressed in meaningful language, meaningful argumentation is always the undeniable objective.

Some philosophers who have studied transcendental argumentation have mentioned, most frequently only in passing, that such argumentation can be put to use in identifying transcendental principles that are necessary for the possibility of meaning.<sup>127</sup> To the author's knowledge, however, transcendental argumentation has not been explicitly used to develop a general theory of meaning. In a later chapter, we shall see how the metalogic of reference underlies the possibility of meaningfulness, in no matter what area of thought or discourse. Here, while we are concerned with the connection between a study of the metalogic of reference and transcendental argumentation, some general comments are appropriate.

We recall from {4} that, for the purposes of the metalogic of reference, the concept of meaning was left open and undefined despite the fact that we shall be concerned to define and validate referential consistency as a necessary precondition of general meaningfulness. Meaning, as we have already observed, may be of many kinds and associated with many things; the broad and inclusive concept of meaning to be understood in these pages is highly general, applying to a wide range of forms of meaning, not restricted to standard propositional or statement-based meaning, but pertaining to whatever can be an object of reference—from abstract, constructed, theoretically hypothesized objects, sets, or formalized representations, to concrete physical objects, to states of affairs, events, perceptions, imagined or fictitious objects, psychological states, dispositions, colors, frequencies, results of calculations, probabilities, relations among any of the foregoing, etc.—in short, whatever identifiable objects comprise the universe in view. The phrase 'meaning as it pertains to an object of reference' is taken here also in a very general sense, so that, for example, even a coordinate point  $(-1, \pi, +\sqrt{2})$  has, by virtue of its

---

<sup>127</sup> Early examples are found in Strawson (1959, p. 35) in which he wrote: "[the skeptic's] doubts are unreal, not simply because they are logically unresolvable doubts, but because they amount to the rejection of the whole conceptual scheme within which alone such doubts make sense." Similarly, Crawford (1961, p. 268) wrote: "I suggest there are principles which are required not merely as methodological principles for determining the truth [sic] or falsehood of certain statements, but as principles which are necessary if it is even to make sense to speak of those statements as being either true or false." Stroud (1968, p. 251) also mentioned in passing "the necessary conditions of anything's making sense." Rehg (1989, p. 257) referred to the "necessary condition for the meaningfulness of the doubt" expressed by the skeptic who wishes to doubt the existence of external objects.



identifiability, a specific meaning for those who can locate that point in an appropriate coordinate system. Such a highly general, framework-relative conception of meaning is likely to be unsatisfying for those who wish for more, but here my wish is only to please the conceptual minimalist. It is a conception of meaning that reflects a recognition of the essential, the transcendental, framework relativity of all identifiable objects of reference. As we shall see in the next chapter, our as yet open and undefined conception of meaning will be described in more precise terms as a function of framework relativity itself.

In applying the criterion of referential consistency, any putative cognitively meaningful statement, proposition, or concept, or any apparent reference to affective meaning, which is such that its referential structure is inconsistent with itself—causing it, as I have expressed this earlier, to undermine or implode upon itself—will be considered, from the reflective standpoint of metalogical analysis, to be devoid of meaning. The generalized criterion of referential consistency as a criterion of meaning will lead to a recognition that this criterion cannot *not* be accepted without entailing metalogical referential inconsistency—ruling out the *possible* meaningfulness of any such rejection.

In {6}, we noted that there is a necessary, inalienable dependency between object identification and the system(s) of reference that provide the basis for such identifiability. The identifiability of a set of objects is inseparable from, and can have no meaning apart from, the reference system that makes reference to them possible. When a concept or proposition is used in a manner that undermines its own capacity in principle to refer, in so doing, as we shall see in detail later, it simultaneously undermines its capacity to be meaningful.

In {7}, we touched on the concept of horizon in connection with the existence of limits beyond which we cannot in principle go in lifting the parameters of constraint that govern possible reference. As we shall explore in detail, when anyone, or any theory or position, seeks to trespass beyond those limits, resulting transgressing assertions will “short-circuit,” becoming referentially incoherent, and thereby meaningless.

The above summary of metalogical results—some, for now, only alleged—has led to the view that parameters of referential constraint delimit the universe of the possible, beyond which lies only unintelligibility, which is to say meaninglessness.

In {8}, we noted that there are three general ways in which presuppositions may be examined: as a function of truth, as a function of meaning, or as a function of identifiability. The inclusive ‘or’ in this statement now stands in need of amplification: When metalogical parameters of reference become our

chosen explicit subject-matter, our interest becomes focused on structural/systemic presuppositions and presuppositions of identification. From the resulting point of view, we recognize that a certain set of preconditions of identifiability is satisfied by the dynamic system in terms of which a specific domain of possible objects of reference is established. From this perspective, presuppositions of truth and those of meaning will be seen to rest on the metalogically more theoretically fundamental presuppositions of identifiability. Possible objects of reference metalogically depend upon a frame of reference that permits their identifiability, while such identifiability provides the necessary basis for both possible meaning and the possible truth-value of propositions or statements about those objects: Without such identifiability, reference to any objects in any context that allows for truth-functional or meaningful thought or assertions is rendered in principle impossible.

Here again it is important to remind the reader that “identifiability” is taken in a maximally general sense, inclusive of both unique, specific-individual identification, and vague, suggestive, even indeterminate or probabilistic identification. Such a conception of identification, though it includes the standard meaning of the term, also extends that meaning by allowing instances in which an object’s identity may only be approximated within limits that are themselves indistinct.

This short retrospective summary of the above observations and claims concerning the general nature of meaning as understood in the metalogic of reference points to the recognition that transcendental preconditions of reference, the central subject-matter of the metalogic of reference, establish *boundaries of possible meaningfulness*, beyond which reference and the identifiability of objects, in principle, are rendered impossible, of whatever kind or nature those objects of reference may be. In a later chapter, we shall begin to call these boundaries ‘*metalogical horizons*’.

In order to investigate these boundaries of possible meaningfulness, we turn now to examine the nature of framework relativity, the subject of the next chapter.

# 10

## Framework Relativity

Only the theory decides what can be observed. [*Erst die Theorie entscheidet, was man beobachten kann.*]

– Albert Einstein, as related by Werner Heisenberg (1969, p. 92)

[T]he object of knowledge is not to be separated from the instrument of knowledge. The Heisenberg principle of uncertainty is one of the expressions of a technical working out of the consequences [of this realization]....

The ultimate instrument is ourselves. This means that not only should we never think of the microscopic world without thinking of microscopes, but we should never think of the microscopic world without thinking of ourselves using the microscope. In general, we should never think of the world around us without also thinking of the nervous machinery in our heads by which we acquire knowledge of the world. To discover the best way of holding ourselves to this awareness constitutes what seems to me to be perhaps our most pressing intellectual problem.

– P. W. Bridgman (1959, pp. 169, 153-154)

**F**ramework-relativity is not relativism, a doctrine which, at the time of this writing, is a pervasive and strongly believed dogma. Framework relativity and relativism are opposed to one another on a theoretically fundamental level. The distinction between the two parallels that between reliable knowledge, the scope of which is defined, and purely provisional knowledge, tentative and uncertain. Relativism, often called ‘cultural’ and sometimes ‘epistemological’, claims—no matter what purported “truths” are discovered or believed to have been proved—that “truths” are no more than provisional, tentative, and dismissible since we have only to remind ourselves that “every-

thing really is subjective,” that “nothing can be known objectively to be true,” etc.

There are, of course, good, strong, and rationally convincing arguments against relativism, the best of which, in my view, points to the fact that the relativist’s own position cannot possibly be articulated without presuming the very objectivity that he or she wishes to deny. But here I am not concerned to refute relativism, only to dismiss it as irrelevant to the concept of framework relativity, which, contrary to relativism, recognizes the objectivity of results relative to the frameworks those results presuppose.

This is not a sophism, not a smokescreen to disguise, elevate, and dignify the “merely relative” by coining an unfamiliar phrase. Let me give some examples of objective results that are framework-relative: Euclidean geometry’s theorems are objective in this sense; they are provably true relative to a set of definitions, postulates, and rules of inference; therein lies their framework relativity. Similarly, Riemannian geometry, which rejects Euclid’s parallel postulate, leads to a set of true propositions, which are provable relative, again, to a specified set of definitions, postulates, and rules of inference. Lobashevskian geometry, which also repudiates the parallel postulate in a different way, leads to demonstrable results relative to an alternative set of definitions, postulates, and rules of inference.

All such results are objectively demonstrable. What meaning does the term ‘objectivity’ have other than simply the capacity to establish results that are not subject to controversy and can be replicated by competent colleagues in a given discipline—*provided* that they accept the framework which the establishment of such results must presuppose? If they don’t accept that framework, they beg the question, they engage in a *petitio principii*—they refuse to abide by the rules of the game at issue, perhaps because they wish to use alternative rules and prefer a different game, or perhaps because they are being intellectually recalcitrant, believing that the skepticism they express is not itself self-undermining (which of course it is). As long as you accept the rules of chess, some moves are *objectively good*, and others are *objectively bad*. As long as you accept the definitions, postulates, and rules of inference of Euclidean geometry, then there are results in Euclidean geometry that are objectively provable. We might call this sense of objectivity ‘framework-relative’, which indeed it is. But we should distinguish such “framework-specific” objectivity from the more theoretically general meaning of framework relativity that applies to the metalogic of reference. There, framework relativity is a concept belonging to the metatheory of reference, a metatheory leading to the recognition that framework relativity is itself objectively provable on a metatheoretical level—by showing that any (possible) attempt to

deny it leads to self-referential inconsistency of a particularly devastating kind.

Mathematicians have long been both aware and appreciative of the essential interrelationship between the formal systems they study and the validity of the results those formal systems entail. Physicists have, during the century since Einstein's special and general theories of relativity, come to understand and also appreciate how the conceptual structures in terms of which their observations both are made and are formulated in the context of a theory reveal an underlying logic of framework relativity. Neither mathematicians nor physicists have, as a result of this growing framework-relative awareness, given up the search for, and claims on behalf of, objectively demonstrable results. This is not because mathematicians and physicists are intellectually stubborn or obtuse, it is because 'objectivity' simply means—and the very possible meaning of the term itself necessarily requires—framework relativity.

It is interesting and importantly relevant to the present study to note that later in his life, Einstein expressed regret that he had chosen the phrase 'relativity theory' (suggested by Max Planck in 1907). He belatedly came to feel that '*Invariantentheorie*' or 'invariance theory' would probably have been more appropriate. In a letter to E. Zschimmer dated September 30, 1921, he wrote:

Now to the name relativity theory. I admit that it is unfortunate, and has given occasion to philosophical misunderstandings.... The description you proposed [*"Invariantentheorie"*] would perhaps be better; but I believe it would cause confusion to change the generally accepted name after all this time.<sup>128</sup>

Had Einstein instead originally chosen the phrase 'invariance theory', he would have coined a name much more appropriately suited to his work, its objectives, and its contribution to physics, a name much less susceptible to the misunderstanding and confusion of "relativity theory" with "the merely relative." Among the invariants that relativity theory has made evident are, for example, the universality of the formulation of the basic equations of electromagnetism irrespective of the choice of reference frame; the invariance of the speed of light in a vacuum for all inertial systems (i.e., not subject to acceleration or gravity); the invariant functional interdependency between energy and mass, and between the curvature of space and gravitation; etc. In all

---

<sup>128</sup> Cf. Holton (2006, p. 269), Nozick (2001, p. 78).

important respects, the special and general theories of relativity seek to formulate invariant, universal laws of nature. Far from exemplifying relativism, Einstein's special and general theories of invariance seek to disclose physical laws that are universal and covariant with respect to the particular, contingent, physical conditions of individual reference frames.<sup>129</sup>

Framework relativity from the standpoint of the metalogic of reference is exactly similar in this respect. Its purpose is to recognize and establish universal principles of invariance that underlie all frames of reference, irrespective of their scopes of interest and sets of possible objects of reference.

### 10.1 Frames of reference, reference frames, and frameworks

[N]o content can be grasped without a formal frame.

– Niels Bohr (1959/1949)

In this work, I use the terms 'framework', 'frame of reference', and 'reference frame' interchangeably. We do not have the physicist's more restricted concept of reference frame in view,<sup>130</sup> but rather that of any system that provides a basis for the identifiability of one or more sets of objects of reference. Frameworks understood in this theoretically general sense are widely used and familiar: A framework insures the possibility that an enterprise has a meaning, a purpose, a set of objectives, and it provides a basis for identifying references to certain sets of possible objects of reference.

In {8.6}, 'identification' and 'identifying reference' were understood in the general, abstract sense in which there is an instance of identification, or of identifying reference, when that which is described is determined as that to which there is reference. We shall make this understanding of identifying reference clearer and more precise as this work evolves; in this broad sense, identifying reference may be understood to involve a descriptive component of specification and a component of satisfaction that is present when what is specified coincides with that to which there is reference. This is made evident when we think of the recognition of an object, of its identity, and realize that

---

<sup>129</sup> For a more detailed and precise discussion, see {26}.

<sup>130</sup> In physics, a reference frame is generally understood as a framework in terms of which observations, measurements, and the mathematical representation of these can be made with the objective of formulating physical laws; such a framework assumes the presence of an observer in possession of one or more time-keeping devices, and the use a coordinate system in relation to which times and positions of events can be assigned. A more specialized physical reference frame is the inertial reference frame of a body, when not acted upon by forces, that either is at rest or moves with a constant linear velocity.

such a recognition is a function of some set of criteria that specify parameters within which that object's identification *can* occur.<sup>131</sup> We are reminded that such broadly understood identification may consist in the unique identification of a particular object of reference, or in a general or vague identification, or consist only in a certain specification according to rule (e.g., as in the *n*th member of a certain series). Hence, the identification of an object of reference may be highly specific, unique to that particular object, or it may be subject to ambiguity, indefiniteness, or probability, and yet, no matter what its degree of specificity, its very identity is, in principle, a function of the frame of reference in terms of which it can be identified. As we saw in {5}, the identity of objects is inseparable, and is logically meaningless, as Quine remarked in passing, apart from the underlying framework, the coordinate system, that permits reference to them.

Identifying reference to an object is only possible when the subject of the identification is in some manner fixed or defined within its presupposed framework in a way that allows for the possibility that re-identification can be made to the same subject. The re-identifiability of objects of reference requires that a frame of reference permit a certain degree of temporally successive, reiterated reference, so that retrospective second-order references are possible to those in the past. A discussion of this temporal basis of identifying reference will be deferred to later discussion.

A frame of reference is that which ultimately supports propositions that can be articulated from that standpoint. On a conceptually most basic level, any statement or question depends for its meaning upon the frame of reference in terms of which it is expressed. In this general sense, "the total meanings of terms are determined by the matrix in which they are embedded" (Bridgman, 1936, p. 5). As we shall see in a later chapter on meaning ({11}), frames of reference and meaning, whether linguistic or non-linguistic, are inextricably related.

In {7.10}, we saw that a frame of reference (framework or reference frame) functions as a theoretically fundamental system of coordination that permits the identifiability of a range of possible objects of reference. Such coordination may, for example, link together spatial locations with objects, times with events, objects and times with observers, etc. The coordinative function of reference frames lies at the very basis of the identifiability of objects, and hence at the very basis of possible knowledge and communicability of such knowledge relating to any class of objects of reference.

It is important that we distinguish, on the one hand, specialized frames of

---

<sup>131</sup> On the logic underlying pattern recognition and its inescapable ambiguity and framework relativity, see Bartlett (2015).

reference, such as are found in individual theories of physics, mathematics, economics, biology, etc., and, in particular, philosophical positions, and, on the other hand, the maximally general frame of reference that is the object of our present study, a framework in terms of which we can study what I shall call the ‘first-order’ structural/systemic preconditions of reference that make possible the identifiability of any specialized framework’s sets of objects. From there, it is a further step of theoretical abstraction to wish to know what kinds of constraints or rules those specialized preconditions of reference themselves obey. The metalogic of reference establishes a maximally general theoretical reference frame in terms of which such first-order constraints can be analyzed with the objective of disclosing the second-order, the metalogical, invariant constraints that apply to any framework of identification whatever. First-order constraints relate to the structural/systemic regulatory mechanisms of individual, specialized frameworks of identification. The framework proper to the metalogic of reference comprises a highly abstract, general identification system whose subject-matter includes all possible frameworks of identification, whereas the first-order constraints are the regulatory mechanisms of individual and special identification-systems.

The second-order, maximally general, properly metalogical framework, because it must be capable of self-applicability (it purports to identify preconditions of all reference, its own preconditions included), is therefore such that the results of the metalogic of reference will render explicit its own transcendental foundation. In this sense, the framework established by our approach might be characterized in Quine’s words (though they were applied by him in an altogether different context): “Our argument is not flatly circular, but something like it. It has the form, figuratively speaking, of a closed curve in space” (Quine, 1963/1953, p. 30). Somewhat later, again in an unrelated context, logician Paul Lorenzen rather similarly wrote: “the method is identified with its own result” (Lorenzen, (1969, p. 89). Reminiscent of Nelson’s and Palmer’s thoughts concerning the peculiar circularity of transcendental arguments, the development of the metalogic of reference must in a sense presuppose that the task is finished before it is begun—i.e., to begin is already to presuppose the *possibility* of the task, which is what that task would seek to demonstrate by formulating its own structural/systemic presuppositions.



## 10.2 Framework-relative field theory

A proposition can determine only one place in logical space: nevertheless the whole of logical space must already be given in it....

– Ludwig Wittgenstein (1961/1921, §3.42)

[E]very partial domain characterizes the total field....

– Hans Reichenbach (1965/1920, p. 103)

[A]ny factor [the essential structure of which relates, necessarily, to the backdrop of fact, apart from which reference it is not itself] necessarily refers to factors of totality other than itself.

– Alfred North Whitehead (1961/1922, p. 308)

In {4.10}, the self-enclosure of a reflective, maximally general theory was described. We noted that philosophical systems are frequently self-enclosed in a similar way, and that for phenomenologists, experience itself comprises a self-enclosed “field” in the sense that no matter how experience is extended, it is never possible to “go beyond” its boundaries, since whatever is experienced becomes part of the field of experience. Later, in {8.3}, we saw how the field concept contributes to a relational understanding of systems, one that abandons a focus on isolated entities and instead studies the interactional principles of organization of totalities.

Here, I want to develop the field concept in greater detail since it sheds light on the nature of frameworks and on the central role that identifiability plays in them. To do this it will be useful to recognize the ways, which parallel those we’ve already described, in which the field concept has been employed in mathematics and in physics.

The basic concept of a field in mathematics was concisely expressed by Richard Dedekind in 1871: “By a field we will mean every infinite system of real or complex numbers so closed in itself and perfect that addition, subtraction, multiplication, and division of any two of these numbers again yields a number of the system.”<sup>132</sup> In this its classical meaning in mathematics, a field is a closed set on which designated arithmetical operations operate so as to

---

<sup>132</sup> Translation from the original German in Kleiner (2007, p. 66).

associate any pair of elements of the set with a corresponding element. In its mathematical sense, a field has three general properties that will be important to us: A field is comprised of a set of elements; certain operations are permitted on those elements; and the application of those operations yields results that are also members of the set, so that the set is, as Dedekind expressed this, “closed in itself.”

In physics, and specifically in quantum field theory, the concept of field is differently conceived: As physics has evolved since Newton, the field concept, introduced in physics by Michael Faraday in 1849, has become increasingly central in the theory of gravitation and electromagnetism, and most especially in quantum theory, leading physicist Steven Weinberg to comment that “quantum field theory has become the most widely accepted conceptual and mathematical framework for attacks on the fundamental problems of physics” (Weinberg, 1977, p. 17).

[M]aterial particles could be understood as the quanta of various fields, in just the same way that the photon is the quantum of the electromagnetic field. There was supposed to be one field for each type of elementary particle. Thus, the inhabitants of the universe were conceived to be a set of fields—an electron field, a proton field, an electromagnetic field—and particles were reduced in status to mere epiphenomena. In its essentials, this point of view has survived to the present day, and forms the central dogma of quantum field theory: the essential reality is a set of fields, subject to the rules of special relativity and quantum mechanics; all else is derived as a consequence of the quantum dynamics of these fields. (Weinberg, 1977, p. 23)

In the years since Weinberg wrote this, fields have continued to serve as the theoretically most fundamental objects of quantum-theoretical study, resulting in the view that the increasingly varied populations of subatomic particles are excitations of and interactions among corresponding fields. Each variety of field is associated with a set of particles of a particular kind.

In its quantum-theoretical sense, a field has two abstract properties that will be important to us: Such a field provides a model in terms of which the elements of a system can be understood—if for a moment we accept Weinberg’s term—as “epiphenomena” of that system, and this model makes it possible to understand those elements in theoretically comprehensive and integrated terms. We need to understand these claims more clearly: Let us

therefore consider an abstract translation of the mathematical and quantum-theoretical concepts of field as they apply to an understanding of framework relativity:

The mathematical concept of field, when abstracted from its application in arithmetical operations on elements of a set, provides a minimalist formal model of what I have termed “self-enclosure”: the permitted operations yield results that remain within the set, which is therefore closed. The quantum-theoretical concept of field, again when abstracted from its concrete subatomic application, provides a conceptual model in terms of which the elements of a system can be understood, in a specific sense that we shall define in what follows, as “epiphenomena” of that system. Framework relativity, as I propose in this study, is characterized by both the self-enclosure of the “referential field” associated with a given reference frame, and by the “epiphenomenal property” possessed by the range of possible objects of reference that are identifiable from the standpoint of that reference frame. Conceptualized in this way, framework relativity provides a way to understand the relational dynamic that intimately ties possible objects of reference to the frameworks in terms of which they are identifiable.

What then does this “epiphenomenal property” mean in terms of identification frameworks? In {5.7}, we saw that reference is essentially tied to identification, that any instance of reference is an instance of identification, whether unique, ambiguous, rule-defined, etc. As an example, we took a specified coordinate point, whose identity is determined as a function of the Cartesian coordinate system that constitutes the basis for its identity and identifiability. Reference to a specific coordinate point presumes an appropriate coordinate system, so that “built into” the identity of such a point are the logically interdependent relations among coordinate system, specification of the point, and the point’s identity.

In {7.3.10}, in the discussion of preconditions of identification, we further saw that frames of reference allow for possible reference to diverse ranges of objects, and thereby make it possible to know what is specified, talked about, described, indicated, etc. The coordinative function of frames of reference was recognized as the basis for our ability to know and to communicate knowledge of any class of objects.

It is then but a small step to recognize that *any object of reference contains, or has ingredient within it, as an integral constituent of its identity, the constitutive structure of the reference frame in terms of which it is identifiable*. The term ‘epiphenomenon’ suggested by Weinberg has been applied with an assortment of different meanings in other contexts and in various disciplines; let us therefore call the fact of this “ingredience” by what may be the

more appropriate and dedicated term ‘*embedding*’.

Recognizing  $(1,0,+\sqrt{2})$  as an identified coordinate point depends upon the fact that the object, the specified point to which there is reference, identified by the symbols ‘ $(1,0,+\sqrt{2})$ ’ brings with it, or “embeds,” the built-in, implicit association of those symbols with an appropriate coordinate system. In a similar fashion, for example, “embedded” in the identity of a perceived object are corresponding logically interdependent relations among an appropriate perceptual coordinate system, the specified perceptual object, and the perceived object’s identity; in the intended sense, the identity of the perceived object embeds these relations.

*Mutatis mutandis*, we may think of many other sorts of objects of reference: a character portrayed in a work of fiction, a mathematical structure identified in a formal system, the taste of papaya, etc. In all these and other cases, the object of reference has ingredient in its very identity, or embeds, the constitutive structure that permits its identifiability. In this sense, the identity of an object of reference is an “instantiation,” a “realization” of the referential field of its presupposed reference frame, in a fashion that parallels a particle’s conceptualization as an excitation of a specific quantum field. This is, as I read their work, what Wittgenstein, Reichenbach, and Whitehead had in mind in the quotations given at the beginning of this section. It is in this sense that “A proposition can determine only one place in logical space: nevertheless the whole of logical space must already be given in it” (Wittgenstein), and that “every partial domain characterizes the total field” (Reichenbach). This property of individual objects of reference—the property of having ingredient within them the overall reference frame in terms of which they are identifiable—is the “epiphenomenal property” mentioned earlier.<sup>133</sup>

It is important that we recognize the distinct levels of reference that have been employed in the foregoing paragraphs. Reference to a perceived object, for example, is not of course reference to the structural/systemic preconditions that provide the basis for its identifiability. The former is object-level reference, and the latter, a second-order, higher, meta-level variety of reference that requires the reference frame proper to the metalogic of reference. In a similar way, specification of a coordinate point, as in the example of  $(1,0,+\sqrt{2})$ , is not itself reference to the overall coordinate system that is implicitly “embedded” in the identity of the specified coordinate point. Here, too, we need to distinguish between the explicit designation of that coordinate point, and the implicit sense in which that point’s identity embeds an appropriate overall coordinate system. We often make a similar distinction clear by

---

<sup>133</sup> In Bartlett (1970, p. 13), this property was called the ‘germinal duplication’ of the overall coordinative system of reference.

talking about the explicit *use* of symbols in contrast to the *reflective* analysis of what those symbols implicitly represent.<sup>134</sup>

### 10.3 Framework relativity and ontology

One's ontology is basic to the conceptual scheme by which he interprets all experiences, even the most commonplace ones.... Judged in another conceptual scheme, an ontological statement which is axiomatic to McX's mind may, with equal immediacy and triviality, be adjudged false.

– Willard V. Quine (1948, p. 29)

It may be alleged that everything exists only from a certain point of view; but (even if it made sense to speak of existing from a point of view) the fact that a certain thing exists from a certain point of view is itself a simple fact, it just is so, and not from any point of view. Even if we could not detach an object from someone's experiencing of it, the whole situation, the person's experiencing of this object, would itself just be there.

– J. L. Mackie (1962, p. 273)

There are clear-cut ontological implications that follow from framework relativity and the role of referential fields. Here, I only sketch some of these and defer a detailed discussion to later chapters.

Once it is recognized that the identity and very identifiability of objects of reference entail relationships of relativity to a reference frame that makes these possible, we are forced to recognize that the very possibility of claims concerning their ontological status are also framework-relative. Metalogical conditions of identification are, as we have seen, preconditions that must be satisfied before an object can be said or thought to have an identity: They are conditions, in this sense, that pertain to formal ontology.<sup>135</sup> Although they underlie conditions of linguistic discourse, they are not themselves linguistic, but rather comprise the logical, formal substructure that makes identification,

---

<sup>134</sup> For a phenomenological study of the explicit-implicit relationship and of conceptual inconsistencies which that relationship often occasions, see Bartlett (1974, 1975a).

<sup>135</sup> See the discussion of presuppositions of identification in {8.13}.

and hence, discourse involving identification, possible. We recall that such metalogical conditions of identification are neither about statements that are true or false, nor about the contextual beliefs held by speakers and hearers. They are, instead, formal, structural/systemic principles that constitute the basis of all identifiability.

The property of the self-enclosure of referential fields itself brings with it ontological implications. As we shall see in some detail later, attempts are often made to grant or to bestow upon objects of reference a degree of ontological independence from the reference frames in terms of which they can be identifyingly referred to; such attempts result in metalogical self-referential inconsistency. This, as we shall also see, leads to the result that it is meaningless *either* to assert the ontological framework-independence of objects of reference, *or* to deny that they have such ontological independence. Both “ontological independence” and its denial are metalogically self-undermining. To be is to be an object of reference, and to be an object of reference is to be an object of reference from the standpoint of an appropriate reference frame. To claim that “to be” is more than this, or to deny this, is to trespass beyond the boundaries of possible reference and meaning.

Ontological matters can quickly become both complex and inflationary due to the great variety of ways in which objects can be referred to in a great variety of frameworks. The framework relativity of ontology cuts through much of this unnecessary complexity by recognizing that any reference frame, relative to the range of possible objects that are identifiable as a function of that framework, establishes some species of ontology, whether purely abstract and theoretical, physically concrete, imagined or fictional, etc. Similarly, identifying reference to any object brings with it, as an “epiphenomenon” of the referential field in which it is identified, necessary relativity to an appropriate frame of reference. Quine’s dictum, that “there is no entity without identity,”<sup>136</sup> finds its ontological basis here.

#### **10.4 Framework self-enclosure and translation to other frameworks**

In previous chapters, I have referred to self-enclosure in connection, for example, with topology, or in connection with many philosophical systems, or in relation to the phenomenologist’s conception of the field of experience, or in connection with the nature of some systems of belief that confine the believer to a point of view that he or she cannot or will not go beyond. But in the present study, our central interest in self-enclosure lies in the metalogical fact

---

<sup>136</sup> See {5.8}.

that it characterizes the reflexive, vertical, non-ordinal reference frame established by the metalogic of reference on the level of maximum theoretical generality. Here, too, but on a metalogical level, the reference frame that is the subject of our study is self-enclosed: We shall find that it is not theoretically possible to “go outside it,” only in part because it comprises a reference frame to study all frameworks of reference, and, reflexively, its own included; the other more important part we shall investigate at some length in subsequent chapters.

However, individual frames of reference—whether they are specialized reference frames such as are used in the sciences, mathematics, and other disciplines, or are the common, everyday frameworks we use in referring to everyday and commonplace objects—these also can possess the property of self-enclosure, but only in a restricted sense: I have employed the concept of a coordinate’s relation to its appropriate coordinate system for the reason that it is conceptually simple and clear. A Cartesian, or rectangular, coordinate system possesses the limited property of self-enclosure in the following sense: As long as the identity of points is specified in the form  $(x,y,z)$ , the points identified lie within a Cartesian coordinate frame. This is of course a truism. Similarly, if coordinate points are specified in polar terms  $(r,\theta)$ , then the points identified must lie within a polar coordinate system. If specified in spherical form  $(\rho,\theta,\phi)$ , then points so specified must lie within a spherical coordinate system; if specified in cylindrical form  $(r,\theta,z)$ , again this is correspondingly true. As long as a specialized system of identification is employed, *and*, for whatever reasons, we are limited to the manner in which that system identifies points, then of course the range of objects of reference which our use of that system can identify has the restricted property of self-enclosure. Although this is tautologically true, we shall see later that, obvious though this fact may seem, it is frequently ignored.

However, for many reference frames it is possible to transform the way in which an object of reference is identified so as to conform to the identification requirements of another system, or systems, of reference, and in that way avoid a reference frame’s self-enclosure. When formal coordinate systems are employed, in a two-dimensional plane we can, for example, transform rectangular coordinates to polar coordinates, or in a three-dimensional volume of space transform rectangular coordinates to spherical coordinates, or rectangular to cylindrical coordinates, and the reverse of these. Formulas taught to high school students convert one system of identification into another.

A similar but more complex example is found in the case where one coordinate frame moves with constant velocity with respect to the other, and we wish to describe the time and location of an event as observed from the stand-

points of the two reference frames. Here it is possible to take the coordinates associated with the event in each reference frame and to perform a translation: For this purpose the Lorentz transformation, named after Dutch physicist Hendrik Lorentz, can be used, incorporated by Einstein in his special theory of relativity. The Lorentz transformation provides a conceptual bridge, a translation, from one coordinate frame to the other, leading to the now-familiar phenomena of time dilation, relativity of simultaneity, and length contraction.<sup>137</sup> These consequences, which come about as a result of interrelating differently constituted reference frames, are, in the context of the present study, significant because they show that not all reference frames can be uniformly translated as is the case in converting a rectangular coordinate point in a plane  $(x,y)$  to its polar form  $(r,\theta)$ . The message to be gained here is that the constitutive structure of a given framework of identification may or may not be directly compatible with—i.e., translatable or convertible into—another. As we shall see later, the unquestioned assumption that fundamentally incompatible frames of reference allow for uniform translation leads to metalogically self-undermining claims.<sup>138</sup>

Translation from one reference frame to another is not always possible when more complex reference frames are involved, without losing the identity of the objects whose identification those reference frames make possible. Attempts to do this, in the belief that the identities of the objects are, so to speak, preserved, frequently run afoul and, metalogically, “short-circuit,” as we shall see later on. Here, it is enough to observe that reference frame translations, or transformations, are sometimes possible, and sometimes not. As we shall see in a moment, transformations between reference frames are possible only when they are, in an important sense, compatible: It would in most cases evidently make no sense, for instance, to attempt to transform a certain set of rectangular coordinates into the reference frame of a novel.

### 10.5 Framework relativity and perspectives

There is nothing improbable in the supposition that an analysis of the world may yield a number of formulae, all consistent with the facts. In physical science different formulae may explain the phenomena equally well.... Why may it not be so with the world? Why may there not be different points of view for surveying it, within each of which all data harmo-

---

<sup>137</sup> More detailed discussion will be found below in {26.6}.

<sup>138</sup> See also Bartlett (1970, pp. 18-22).



nize, and which the observer may therefore either choose between, or simply cumulate one upon another? A Beethoven string quartet is truly, as some one has said, a scraping of horses' tails on cats' bowels, and may be exhaustively described in such terms; but the application of this description in no way precludes the simultaneous applicability of an entirely different description.

– William James (1905/1897, p. 76)

In {8.6}, I listed five of the common varieties of identification: unique, rule-based, general, abstract, and vague. We saw that frameworks provide the basis for identifiability, and advanced the preliminary claim that framework relativity brings with it the relativity of ontology in the sense that to be an object is to be an identifiable object of reference relative to an appropriate reference frame. We further saw that embedded in the identity of an object of reference is what I've termed its 'constitutive structure'—which is to say that functionally constitutive of an object's identity are the logically interdependent relations that underlie the identifiability of that object relative to its appropriate frame of reference. In the above senses, to every constituent of the universe is associated a form of identity, and hence to which is associated one or more reference frames in terms of which it is identifiable.

We saw in the previous section that, for many reference frames it may be possible to transform the way in which an object of reference is identified so as to conform to the identification requirements of another reference system or systems, and in that way avoid a reference frame's restricted sense of self-enclosure. Rules for converting standard systems of coordinates from one kind of coordinate system to another, and the Lorentz transformation used in special relativity, served as examples. And yet, as we shall see, such transformations, translations, or conversions are not always possible.

We frequently make recourse to the notion that objects of reference—whether they are physical, abstract, or fictitious objects, or events, propositions, theories, or concepts, (etc.)—may be considered from multiple *perspectives*. This commonly used notion in its application to reference frames is important and can profit from some reflective analysis.

By a “perspective” what I will mean is any one out of a group of compatible reference frames in terms of which a given object or set of objects is identifiable. What then are “compatible reference frames”? We shall mean any frames of reference in terms of which it is possible to refer to *the same object*. To give an example, again in terms of coordinate systems: If (5, 25) specifies

a coordinate point in a Cartesian coordinate system, then (25.495, 78.690°) specifies the same point in polar coordinates.<sup>139</sup> What does “the same point” here mean? It means that there exist rules that enable us to transform the rectangular coordinates into corresponding, equivalent, polar coordinates so that, for every point in one coordinate system, an identical corresponding point in the other coordinate system can be specified.<sup>140</sup> In short, such rules establish an method of translation and identification, so that corresponding to the point specified in rectangular terms we can identify a point specified in polar terms such that the two points coincide.

The same meaning of compatibility carries over when we consider the compatibility of reference frames that are not formally rule-governed in the precise way in which coordinate systems are: In general, for any two reference frames to be compatible, each must, first, provide a means for translating the identity of an object given in that frame into a correspondingly identified object in the other frame, and second, the translation must accomplish this in a way that preserves the identity of the object, so that it is justifiable to claim that the two reference frames identify the same object.

Consider the two distinguishable frameworks exemplified by, on the one hand, human visual perception, and, on the other, human tactile perception. Suppose that (1) a stick is placed so it is partially immersed in a glass of water and allowed to remain there; (2) we see in the glass of water a stick that is bent; and (3) we touch in the glass of water a stick that is straight. Suppose further that we’re able to associate particular bumps on the bent stick that we see with corresponding particular bumps on the straight stick that we touch. We are, furthermore, similarly able to recognize that the seen glass-and-the-water-it-contains corresponds to the touched glass-containing-water. We conclude (of course this usually happens automatically, without intervening reasoning—unless we’re philosophers) that there is but one stick, which, when we’ve drawn it out of the water, we both see and feel that it is straight, and therefore we retrospectively conclude (again, this is usually automatic) that it was straight all the time, and that the theory of refraction explains why, from the standpoint of visual perception, it appeared to be bent.

Let us look at this situation a little more carefully: Notice that we have established a *third* framework, a meta-framework, that allows us to refer to the two reference frames, one of vision, one of tactile sensation. From this vantage point, we are able to interrelate these two, now-subordinate, reference frames. Being able, for example, to correlate visual bumps with corresponding

---

<sup>139</sup> The number of degrees is given only to the thousandths place.

<sup>140</sup> The two rules are: to transform rectangular  $(x,y)$  to polar  $(r,\theta)$ ,  $\sqrt{(x^2 + y^2)} = r$  and  $\arctan (y/x) = \theta$ .

tactile bumps would allow a primitive artificial intelligence device to recognize that a single object is in view and is being touched. —But notice that here, a definition of “single object” has to be presumed by the device: A “single, self-same object” is defined for the machine by the correspondences (e.g., bump<sub>a</sub>-to-bump<sub>a'</sub>, etc.) which the two subordinate reference frames identify.

From a purely theoretical, and ignoring the practical, point of view, there is nothing that compels us to make this presumption of a single unifying object, and that presumption may in some instances be questioned. In fact, the unifying presumption that insists on there being one single underlying object of reference is not, for example, made without qualification in the quantum domain, as when two different experimental setups, and the identification frameworks they presuppose, provide evidence that light, on the one hand, is a particle phenomenon, and, on the other hand, that it is a wave phenomenon (see {27} for further analysis).

Two central concepts emerge from our simple examples: These are the concepts of *isomorphism* and *complementarity*. Two frameworks of reference are, from the standpoint of the metalogic of reference, *isomorphic* if and only if the constitutive structures of the two frameworks are such that any object identified in one framework can be placed in a one-to-one relation with an object in the other framework, and in a manner that satisfies criteria that qualify the two objects as being identical, i.e., one and the same object. We recall that when we speak of the “constitutive structure” of a framework of reference we have in view the structural/systemic presuppositions that must be satisfied in order for reference to objects in the given framework to be possible. In other words, isomorphism here does not have the more common meaning of simply “possessing the same structure,” as might be the case when comparing the structures of two wooden crates, but it rather has the meaning of possessing the same structure on the modal/transcendental level studied by the metalogic of reference: It is, we might appropriately call it, ‘metalogical isomorphism’. From a meta-standpoint, two reference frames are metalogically isomorphic if the positive and negative parameters of constraint of both frames respectively determine the same sets of possible objects of reference. We shall come to the second concept, framework complementarity, in a moment.

Earlier in this section, I mentioned that transformations or translations between dissimilar reference frames may not always be possible. To make this clear with the simplicity that coordinate systems provide, let us suppose that  $(a_1, b_1, c_1)$  are rectangular coordinates of reference frame A that specify one vertex of a pyramid, and that three other rectangular coordinates — $(a_2, b_2, c_2)$ ,  $(a_3, b_3, c_3)$ ,  $(a_4, b_4, c_4)$ —similarly specify the pyramid’s other three

vertices. Reference frame A makes it possible to describe volumes, and in the case of the pyramid, its volume.

Now consider a two-dimensional rectangular coordinate system, B, whose coordinates have the form  $(x, y)$ . B makes it possible to specify coordinates of points that determine areas in a plane.

The question we now wish to ask is whether the identity of objects, such as the pyramid, which are specified in A, a reference frame which provides a context in terms of which the concept of volume has meaning, can be transformed into identical objects in frame B, and still allow the meaningful use of the concept of volume.

The fact that A and B are *dimensionally incompatible*—A allows for reference to 3-dimensional objects, and B permits reference to those that are 2-dimensional—leads to the evident fact that the concept of volumes does not retain its meaning in B. (We could, as in projective geometry, obtain projections in B of the pyramid in A, but we can no longer speak of “volumes” in B while preserving the same meaning of ‘volumes’ that A makes possible.) An object’s volume as given in A cannot—without loss of the identity of that object, particularly here its property of having a volume as understood in A—be translated into reference frame B.

An object of reference may be said to have a *dimension* in the general mathematical sense—i.e., its dimension is the minimum number of coordinates needed to specify it within the reference frame that makes reference to that object possible. The notion of dimension is clear-cut in connection with coordinate systems: As we’ve already seen, two coordinates are needed to specify a point in a 2-dimensional rectangular coordinate system; three are needed in a 3-dimensional system; etc. We tend to associate dimensions with space and time, but many other sorts of dimensions may be needed to specify the identity of an object. Any property, relation, or factor that is shared, to varying degrees, or perhaps not shared at all, by the objects in a set may function as a dimension that makes possible the identifiability of those objects. The heights, weights, hair color, ethnic background, language proficiency, IQ, academic standing, career interests, etc., may all be employed as dimensions that make it possible to identify, for example, individual members of a football team, and numerical values for each of these dimensions could be assigned, which would then function as coordinates that identify team members. As the number of dimensions increases and they are measured with increasing accuracy, individual team members can be uniquely identified by the specified coordinates.

Earlier in this chapter, we saw that any object of reference contains, or has ingredient within it, as an integral constituent of its identity, the constitu-

tive structure of the reference frame in terms of which it is identifiable. From the very concrete level of an object's physical dimension (again, the minimal number of coordinates necessary to specify its identity), we see this expressed in the above example: Each object (AKA each team member) has, to a greater or lesser extent, as an integral constituent of his or her identity, the dimensions which, in our reference frame, we've specified in terms of which to identify them. The fact that these dimensions are, so-to-speak, "built-in," "embedded," or "intrinsic" in the coordinates that represent measured values of those dimensions, is straight-forward and free of mystery.<sup>141</sup> It is no different with respect to the constitutive structure of an object of reference; it is merely theoretically more abstract.

The dimensional incompatibility of reference frames should now take on a clear meaning. Such incompatibility is one way in which translatability from one reference frame to another—without the loss of the identity of objects of reference and loss of meaning associated with their identity—is rendered impossible. And there are also other ways, as we shall see later on.

To return now to the concept of framework complementarity mentioned earlier: Two frameworks of reference that are dimensionally incompatible will here be considered *complementary* when, from the standpoint of a third (meta-) framework, the set of objects identifiable in one frame qualifies as the same set of objects identifiable in the other frame. In our uncomplicated example of a stick in a glass of water, the theory of refraction plays the role of this meta-framework, providing a basis for qualifying the visually perceived and the touched objects as being the same.

We clearly select and apply different sets of criteria in determining under what conditions objects given in different reference frames are to qualify as the same. The reference frames in question can sometimes be fundamentally very dissimilar, as in the case of the different reference frames of quantum experiments that imply contrary properties of light, and yet meta-level criteria, for whose acceptance there may be theoretically compelling reasons, may lead us to accept that objects given in reference frame<sub>1</sub> qualify as the same objects from the standpoint of reference frame<sub>2</sub>. Under these conditions, we shall call frame<sub>1</sub> and frame<sub>2</sub> 'complementary', and call the descriptions of objects to which those frames permit reference 'complementary des-

---

<sup>141</sup> A parallel example is found in the non-philosophical area, the study of signal processing, indispensable in the development of computer vision. Here, the number of variables required to describe a signal define its so-called "intrinsic dimension." Another parallel example comes from mathematics in which the dimension of an object is considered to be an "intrinsic" property, independent of the logical space in which the object may be given.

criptions'.<sup>142</sup>

We can now define 'perspectives' more clearly using the concepts we have introduced: When two sets of objects of reference, given in two different reference frames, are considered to be either "the same" or "complementary," then those reference frames establish two different perspectives. The concept of perspectives will later come to be useful when we analyze specific examples in which reference goes awry due to a lack of understanding, neglect, or denial of the structural/systemic presuppositions of reference of individual frameworks of reference.

### 10.6 Framework relativity and conceptual constructs

It will be evident to most readers that some of the concepts I have introduced in this chapter are conceptual constructs designed to help formulate and to understand the nature of systems that make it possible to identify different ranges and kinds of objects of reference. The quantum-theoretical fields mentioned earlier in this chapter possess a reality as objects of reference that is *a function of* the mathematics that quantum theorists have so far found to be necessary to account for the microphysical phenomena they observe. We are not normally accustomed to recognize that "the reality" of an object is a function of the framework of identification that is required to identify it, but it is hoped that readers of this study will gradually become accustomed to thinking about and to understanding "reality" in that way. For, as we shall see, to try conceptualize reality and understand it in any other way leads to inconsistencies that ultimately *can have* no sense.

Although "referential field" and the "embedment" of the overall framework in the individual object of reference are evidently theoretically abstract concepts, from the meta-level, reflective standpoint we are developing, the concept of referential fields and of embedment of the overall reference frame will allow us to recognize more clearly certain of the inevitable boundaries of reference and of meaning.

---

<sup>142</sup> For an informal study of the relativity of facts in relation to the frameworks of reference in terms of which those facts are established, and the resulting understanding of such facts in terms of complementarity, see Bartlett (1975b).

# 11

## The Metalogic of Meaning

It is appropriate to begin this chapter with the occasional reminder of the definitions of two central terms used in this book. As discussed in {4.8}, by the term ‘*metalogic*’ in this study is meant a methodical, logical study undertaken from a superordinate frame of reference—that is, from a level of maximum theoretical generality. By the phrase ‘*metalogic of reference*’ is meant a study from the level of maximum theoretical generality of the general principles that govern possible reference in any subordinate frame of reference, as well as inclusive of the metalogic itself. From this reflexive meta-standpoint, the phrase ‘*metalogic of reference*’ is used in this book specifically to refer to the maximally general investigation of the preconditions of identification/identifying reference entailed by the reference frames of subordinate concepts, theories, or positions, as well as entailed by its own superordinate frame of reference.

Hence, this chapter’s title should lead the reader to expect that we shall begin a meta-level study of certain preconditions of this kind as are entailed by what we understand by “meaning.”

...

The principal objective of this study is to investigate the limits of meaning, the limits of that about which we can think and talk meaningfully. But not merely the limits of meaning of these things, but the limits of their *possible* meaning. To this end, we shall of course need to understand some fundamental things about the nature of meaning, and, more specifically, about what it is that makes it *possible* for anything to have a meaning, or to be meaningful.

There are two basic ways we could hope to do this. One is to seek to understand, at a theoretically fundamental level, what must be granted in order for *meaning of any variety* to be possible. This would provide us with a description of a *necessary condition* of possible meaningfulness.

Another way would be to consider *theories of meaning* as they have been developed by philosophers, and with reference to those theories as a group, seek also to discover that which must of necessity be granted in order for such

theories and their subject-matters to be capable of possessing whatever meaning they propose to have.

We shall use both approaches in this chapter, each designed to advance the central interest of this work: to determine limits of possible meaningfulness—boundaries beyond which we cannot, in principle, meaningfully transgress.

To accomplish this, we shall, in turn, also need to understand something about the nature of meaninglessness; we shall then be concerned to identify and to describe a fundamental *sufficient condition* of meaninglessness.

### 11.1 Meaning and theories of meaning

Meaning is ... just the sort of word with which we may attempt to probe the obscure depths of the souls of fishes. “Let us fix attention on the state of the mind of the goldfish.... Suddenly comes a new element into consciousness—the conscious counterpart of the stimuli of the eye caused by the bread falling into the water. The food is an object in space and time for the fish and has its MEANING, but when the food is eaten both percept and MEANING disappear.... This is an instance of percept and MEANING tied.”

– C. K. Ogden and I. A. Richards (1923, p. 179,  
quoting W. E. Urwick, 1907, p. 68)

The obscure depths of the souls of fishes is a suitable place to begin a discussion of theories of meaning, for they are as numerous as the fish of the sea and just as slippery. The now little-read, old classic *The Meaning of Meaning* by C. K. Ogden and I. A. Richards (1923) contains a chapter, “The Meanings of Philosophers,” which bulges at the seams with hilarious quotations expressing the astounding and at times perverse multiplicity of ways in which philosophers have theorized about meaning. I will not propose to lead us on a long detour to describe the numerous adventuresome paths that have been taken. Instead, in the following table is my own much-abbreviated list of just a few of the many theories philosophers have proposed to understand meaning.<sup>143</sup>

---

<sup>143</sup> It would require a long, detailed, and complex expository digression to associate the names and works of individual philosophers with the theories of meaning which they have at different times developed. In many cases, a philosopher will incorporate aspects of more than one of the theories I distinguish, and then will subsequently change his or her views. To disentangle the convoluted history of philosophical theories of meaning is not my purpose here, but rather to



(1) *Concept-based or mental imagery theories*: The meaning of a word, phrase, or sentence is equated with an idea, a thought, a proposition, or mental imagery with which it is associated.

(2) *Activity-based, psychologically based theories*: Meaning is understood to be the result of purposeful volitional activity on the part of a speaker: He or she actively intends to communicate a certain message to a listener; if the speaker is successful in achieving his or her intention, and the listener satisfies conditions of a receptive listener, the meaning of that message is communicated to the listener.

(3) *Contextual-relational theories*: Meaning is understood in relational terms as a function of a speaker, a hearer, and the context in which communication occurs between them.

(4) *Referential theories*: Meaning is viewed as the relation between a symbol and what the symbol stands for, or meaning is equated simply with that to which the symbol refers.

(5) *Pragmatical theories*: The customary, authorized, public use of a word, phrase, or sentence by a group of people determines its meaning.

(6) *Rule-based theories*: Rules define the meanings of symbols and their combinations in a natural or artificial language; such rules define permitted vs. prohibited uses of those symbols and combinations, often subject to contextually defined conditions.

(7) *Dictionary-based theories*: The meanings of words and combinations of words are defined in terms of synonymy with corresponding meanings of words and combinations of words that have acquired established identical, similar, or overlapping meanings.

---

distinguish certain of the major theories of meaning and then to describe the range of forms of meaning they have been designed to study.

Readers unfamiliar with philosophical theories of meaning may find the following sources useful as a springboard to the extensive literature: Beaney (1997), Bloomfield (1933), Carnap (1942, 1947), Chomsky (2000), Davidson & Harman (1972), Davis (2002), Evans & McDowell (1976), Horwich (2005), Kripke (1972), Larson & Segal (1995), Morris (1946), Ogden & Richards (1923), Preyer & Peter (2005), Putnam (1975), Quine (1960), Russell (1903, 1940), Schiffer (1972), Soames (2003, 2010), Toribio & Clark (1998), Wittgenstein (1961/1921, 1953).

(8) *Information-based theories*: Meaning is understood in information-theoretic terms: The meaning of a message is its information content, which may be quantified, e.g., in terms of the degree to which the message reduces uncertainty or entropy, or increases the probability of certain outcomes.

**Table A**

Theories of meaning may be considerably restricted in their scope of inclusion or application. For instance, the dominant interest today among Anglo-American philosophers is the semantics of natural language, with emphasis upon speakers' and hearers' linguistic activity in relation to context, belief, and as a function of communally accepted use. Linguistic meaning, however, comprises only a small part of the spectrum of meaning. The following is but a partial listing of major varieties of meaning; as the reader will note, linguistic meaning is represented by only some of the voices in the large choir.

Meaning is, or is expressed by:

- a. a relation of synonymy in dictionaries as found among words and expressions claimed to have similar or equivalent applications, uses, designations, etc.
- b. variously understood semantical associations of a sentence with a proposition, truth-value, or an extensionally defined object or set of objects
- c. the logical, behavioral, or psychological consequences of a state of belief
- d. the set of ways in which a given statement can be verified or shown false, or otherwise known to be true or known to be false
- e. the set of operations in terms of which a term, concept, statement, or theory is employed
- f. the idea or mental image summoned up by a word, phrase, sentence, or other symbol

- g. a state of emotion, memory, or other psychological state associated with an event, object, person or persons, symbol, work of art, experience of nature, etc.
- h. the physical consequences of an action or process
- i. logical possibility; that which is not prohibited by the rules of a formal system
- j. the logical or anticipated consequences of an argument; position; theory; law, legal contract, precedent, or claim; etc.
- k. the place or role of a statement, axiom, postulate, etc., within a system formed by these and other statements, axioms, postulates, etc.
- l. the role played or the objective served by a move or group of moves in a game
- m. the anticipated implications of a person's, a group's, or a non-human animal's behavior
- n. an interpretation of a work of fine art, music, literature, etc.
- o. the interpretation of texts, generally
- p. the role of dreams, prophesies, prayers, other subjective states, etc.
- q. an individual's set of life-objectives, values, commitments, etc.
- r. a view of the world, a *Weltanschauung*, in terms of which a person situates and understands his or her life, work, efforts, etc.

**Table B**

The eight common varieties of philosophical theories of meaning in Table A have occupied themselves primarily with a study of the kinds of meaning that appear toward the top of Table B: Generally, at the time of this writing, philosophical theories of meaning tend to concentrate on a. – d. Meaning as defined by e. and f. have fallen out of most philosophers' scope of active interest, while g. – r. have not often found a place among the central concerns

of the philosophical theories of meaning that I've listed. The study of some of these latter kinds and sources of meaning is today often delegated to other disciplines—for example, psychology, physics, the theory of formal systems, law, game theory, literary and art criticism, religion, etc. Nonetheless, despite changing philosophical fashions, tastes, and dominant paradigms, none of the varieties of meaning a. – r. can be said to comprise an *implicitly inappropriate* subject for philosophical study.

It should immediately be evident from Table B that philosophical theories of meaning have generally taken as their preferred subject-matter only a subset—largely linguistic in nature—of the extensive variety of forms of meaning with which we are all familiar. This observation need not be taken as a criticism, but it serves as a reminder that the range of meaning is considerably larger than that examined by mainstream philosophy today.

As we consider the collection of eight philosophical theories of meaning in Table A, we recognize that all comprise frameworks of reference in terms of which the forms of meaning in which they are interested can be identified. Some of these theories qualify as *perspectives* and some may constitute *dimensionally incompatible* reference frames, as these terms were defined in the previous chapter. We recall that two reference frames comprise perspectives when they are compatible in the sense that they permit reference to the same set of objects. Theories (2)–(7) (activity-based, contextual-relational, referential, pragmatic, and rule-based theories) can be grouped together when each approach offers a theoretical perspective concerning a shared set of *dimensions* of linguistic meaning—i.e., the same set of linguistic factors, properties, relations, etc., which are regarded by those theories as constitutive of linguistic meaning. Often, a given philosopher of language will integrate certain features of some or all of these six theoretical perspectives concerning the nature of linguistic meaning. Theories of the first kind—concept-based/mental imagery theories—may be dimensionally incompatible with theories (2)–(8) when, as is often the case, a philosopher of language wishes to avoid anything resembling a “psychologistic,” mental concept/imagery understanding of meaning. The last approach on our list, information-based theories, has yet to be integrated into current prevailing philosophical discussion concerning the nature of meaning.

But whether a theory of meaning is compatible or incompatible with others, *all* theories of meaning seek to formulate frameworks of reference in terms of which the forms of meaning they analyze *can be identified*. Here lies their interest from the standpoint of the metalogic of reference. No matter what subset of the numerous varieties of meaning (Table B: a. – r.) a theory of meaning studies, and no matter how it proposes to undertake its own approach

(Table A: (1)–(8)) in studying that subset, it is evident that it cannot, in principle, make a beginning unless its reference frame provides for the possibility of identifying reference to the subject-matter—the particular form or forms of meaning—that it wishes to study.

In {4.9}, the definition of meaning was left open. In light of the extreme diversity of kinds of meaning that we see in Table B, it should be clear why a unitary definition that could encompass all forms of meaning has not been offered. Instead, *a necessary precondition of meaning* in the form of *referential consistency* will be formulated; we shall define that precondition in a few moments.

We may recall that the relation between the identification of a certain set of objects and the system of reference that makes this possible is one of necessity: A reference frame provides a basis for identifying reference to the set of objects whose identities that framework can specify—whether uniquely, vaguely, or in a rule-determined manner. There is an inevitable, metatheoretically necessary, relativity of the identity of objects and the frame of reference permitting their identification. As observed in {6}, the identity of a set of objects of reference *can have* no meaning apart from the reference frame that makes identifying reference to them possible. Let us call the claim made in the preceding sentence, ‘claim *P*’. In that sentence the term ‘meaning’ is used in the comprehensive sense that encompasses whatever forms of meaning are—or *might* be—included in Table B. Any additional variety of meaning not already included in Table B *would need to be identified* in order to add it to the list in Table B, and for such identification to be *possible*, an appropriate reference frame must of necessity be presupposed. This is important to see. By virtue of this metalogically inescapable fact, claim *P* gains renewed validation. We *cannot*, in principle, add any new variety of meaning to Table B without identifying it. If we could, this would contradict claim *P*. Therefore, our decision to leave the concept of meaning open has done its job by recognizing that, despite the decision to leave ‘meaning’ undefined, we have nonetheless specified a necessary condition of meaning: Meaning, as well as any theory concerning it, necessarily requires identifiability.

## 11.2 Referential consistency as a criterion of meaning<sup>144</sup>

Criteria of meaning proposed by philosophers in the past have failed to persuade general acceptance. I use the word ‘persuade’ here intentionally, since,

---

<sup>144</sup> This portion of the chapter is a descendant of Bartlett (1982), with a number of corrections, changes, and additions.

other than attempts to *persuade* others to accept the merits of a given criterion of meaning, no criterion that has so far been proposed *compels assent to it*. It has always been possible for a philosopher to choose *not* to accept a given criterion of meaning without thereby becoming *incoherent*—without, that is, giving up the very possibility that *rejecting* that criterion of meaning *can make sense*. As a result, criteria of meaning recommended in the past have often reflected prevailing scientific practice, or have simply argued for the adoption of strict empirical or logical standards of justification. The historical failure of criteria of meaning to gain universal acceptance has been due to their apparently arbitrary—i.e., non-necessary—status as standards *external* to the sets of statements to which they would apply. Often, such criteria have also failed to qualify as meaningful in the test of their own self-application.

My intent here is to show that there exists a metalogically inescapable criterion of meaning that *must* be satisfied in order for individual claims, concepts, and frameworks to qualify—in *any* sense whatever—as “meaningful.” I’ve called that criterion ‘referential consistency’. It is a criterion of meaning in the largely negative sense that non-satisfaction of the criterion involves a certain type of meaninglessness whose investigation was initiated and then studied by the author over a period of decades. As mentioned in the preceding section, this criterion does not aim to formulate a sufficient condition of meaningfulness, and, in light of the extensive varieties of forms of meaning we’ve already encountered in Table B, one may indeed seriously doubt whether a sufficient, all-embracing condition *could* be formulated. And yet the proposed criterion defines an important *lower limit* of meaning, below which claims, concepts, and frameworks become self-undermining and incoherent. It is in this latter sense that we shall find that this criterion can serve as a compelling conceptual tool for the internal analysis and criticism of specific claims, frames of reference, philosophical positions, and scientific theories.

The criterion I shall suggest has these unique properties: Acceptance of the criterion is non-arbitrary—which is to say, it is compelling in a sense we shall explore in a moment. And applications of the criterion avoid begging the question in a way in which appeals to external standards do not.

Logical criteria for evaluating, e.g., the validity of an argument or for assessing the consistency of a theory define standards or limits of acceptability which such argumentation or theory construction presupposes or is thought to presuppose. To a large degree, such criteria are “arbitrary” in the sense that they can be changed if our purposes are served by such a change. Seen as conventions that we accept in the light of our objectives,<sup>145</sup> the criteria that

---

<sup>145</sup> For an early recognition that logical rules may be understood essentially to comprise conventions that we agree upon, see Curry (1957). See also below, note 152.

delimit what we will accept are seldom absolute. That is, we are seldom compelled, *on pain of incoherence*, to accept certain specific criteria rather than certain others, although it is often the case that, if we are to hold to our purposes, we must abide by those or related criteria if we are to accomplish what we intend.<sup>146</sup> In general, then, I shall call a criterion ‘*weakly non-arbitrary*’ or ‘*compelling in the weak sense*’ if non-satisfaction of that criterion precludes achieving the task at hand, and ‘*strongly non-arbitrary*’ or ‘*compelling in the strong sense*’ if non-satisfaction of the criterion results in meaninglessness. These two senses will become increasingly clear as our discussion progresses.

The criteria that define what we mean, e.g., by ‘validity’ and ‘consistency’ are “logically arbitrary” in several ways: If we detect that a criterion or—equivalently here—a rule has been broken, we are free to amend the rule (and perhaps in so doing change the ends which the rule may serve), or correct the violation, or leave things as they are, or shift our perspective, perhaps to a more general point of view and perceive the breaking of the rule as conforming to a broader rule in relation to which it is no longer identified as a violation. And we may have other options. But whatever the special nature of the case may be, criteria of the sort used to assess the validity of arguments and the consistency of theories constitute what may be viewed as logically arbitrary rules for playing certain games: Such rules are the logical features of specific practical or theoretical activities; the control which they make possible is a control that we choose to have, and we are at liberty to choose otherwise.

In relation to the particular purposes that we may have, a certain set of logical or other evaluative criteria rarely compels us by reason of logic alone to accept those criteria and no others. There is, often and in general, a sense of “open-texture” about our objectives. The formal constraints we accept may be selected because they reinforce other ends we intend: economy, comprehension, concinnity, etc. *How* we do or *should* make selections from among alternative, logically arbitrary criteria will not be examined here.

From the standpoint of the criteria we accept, our purposes are underdetermined or specified with a degree of vagueness to just the degree that these criteria are logically arbitrary. It is fair to say that attempts to delimit meaning by means of any particular proposed necessary and sufficient criterion of meaning have failed precisely because of this logical arbitrariness. The numerous criteria that have been recommended for detecting meaningless concepts and statements have very much the same status as do criteria which permit evaluations of validity, consistency, etc. Criteria of meaning that have

---

<sup>146</sup> Relevant here is Wittgenstein’s attention to the relationship between using rules and achieving practical ends. See, e.g., Wittgenstein (1956, I- 9, 20, 131, 162; V- 31ff; and *passim*).

so far been proposed may be considered in the same game-relative light as rules of logical evaluation.

To give several examples: Hume, Schlick, Ayer, and Carnap have proposed the following criteria of meaning:

For Hume: expression of abstract or empirical reasoning.<sup>147</sup>

For Schlick: association of conditions with a proposition or question which define what experience(s) would make that proposition true, or which would if satisfied answer that question.<sup>148</sup>

For Ayer: verifiability, reflecting an individual's knowing how to verify a proposition that is factually significant to him.<sup>149</sup>

For Carnap: ability to give rules according to which observable effects can be deduced,<sup>150</sup> or alternately, expression of factual content.<sup>151</sup>

These criteria, not exhaustive of those that have been proposed, nor mutually exclusive, share two characteristics: First, acceptance of one or more of these criteria is a function of one's purposes, of the objectives of a theory or system, etc. Second, neither Hume, nor Schlick, nor Ayer, nor Carnap, nor any other proponent of a criterion of meaning has yet been able to show that acceptance of a certain criterion of meaning strongly compels assent, i.e., is non-arbitrary in the strong sense.

This observation would not reflect a negative judgment if, as might be claimed, we only wish a criterion of meaning to function with the same measure of arbitrariness in the framework of a set of concerns as does a rule-based convention of logical evaluation.<sup>152</sup> But this state of affairs would clearly not satisfy many of those who urge the adoption of meaning criteria.

<sup>147</sup> Hume (2004/1748, sec. XII, iii).

<sup>148</sup> Schlick (1959/1932-33).

<sup>149</sup> Ayer (1952/1936, p. 35).

<sup>150</sup> Carnap (1935, pp. 13-14).

<sup>151</sup> Carnap (1967/1928, pp. 325ff and passim).

<sup>152</sup> Such arbitrariness in the domain of theory often does not carry over to human practice; there is, to be sure, often a strong sociological-psychological component that enters in: Certain logical rules are "hard-programmed" in society so that their *rejection* is counterintuitive for the majority of people, as, for example, when non-distributive lattices or suspension of the law of excluded middle are employed in quantum theory, leading to interpretations that conflict with habitual ways of thinking and which therefore are felt to be unnatural and objectionable.

The matter is the other way around when it comes to criteria of meaning, since *violations* of such criteria tend heavily to populate the domains of ordinary, and of some technical, dis-



We recognize then that traditional criteria of meaning function in an *external* capacity: When they are applied, they are used to evaluate statements, concepts, or frameworks, as it were, from the outside. Criteria of meaning, understood as stipulative, normative conventions, can only be recommended and argued for in a manner that seeks to *persuade* our acceptance—to convince us—since they do not, in and of themselves, strongly compel assent.<sup>153</sup>

One of the most persuasive cases that can be made on behalf of the choice of a certain criterion of meaning is that its meaningfulness follows from its self-application.<sup>154</sup> If a particular criterion recommends, for example, that meaning be identified with expression of factual content, it may be argued that the very expression ‘factual content’, understood in terms of operations that define the criterion, itself expresses factual content.

However, the self-applicability of a criterion of meaning, even when such application is self-consistent, at most insulates the use of the criterion from internal inconsistency, and may strengthen the *feeling* that its choice is not arbitrary. Beyond this, self-applicability does not do much: The decision to adopt a particular traditional criterion of meaning remains external to the class(es) of statements and concepts to which it is to apply.

### 11.3 Referential consistency as an intrinsically determined criterion of meaning

We have noted how rules for evaluating logical validity and consistency as well as criteria of meaning share the property of arbitrariness as game-relative conventions, and how such rules and criteria gain acceptance as a function of sets of objectives. We’ve also made the observation that criteria of meaning proposed in the past offer little more to persuade one to accept them other than simply the fact that they may be consistent with a set of chosen objectives and hence with conventions that facilitate the realization of those purposes. In any given field of study, rule-based evaluative conventions of one kind or another may undeniably be convenient, expedient, or necessary in practice; if one chooses to work in that field, it can be helpful or even neces-

---

course. As a result, proposals that criteria of meaning ought strictly to be adhered to also result, but for a very different reason, in counterintuitive reactions among many people, and can be resisted with great rigidity and stubbornness.

<sup>153</sup> Carnap’s introductory sentences in his *Logical Structure of the World* come to mind: “What is the purpose of a scientific book? It is meant to *convince* the reader of the validity of the thoughts which it presents” (Carnap 1967a/1928, p. xv, italics added).

<sup>154</sup> Richman (1953) and Schmidt (1957) may be mentioned as among the first philosophers to discuss the importance that a criterion of meaning be self-referentially meaningful.

sary to be able to make recourse to a set of externally imposed evaluative conventions. But the use of such external standards of evaluation cannot, as we have seen, be expected to be non-arbitrary and compelling in the strong sense.<sup>155</sup>

In contrast to criteria of meaning that have been proposed in the past, *there does exist a metalogically compelling basis for evaluation, a basis which one cannot not accept*, which I've called 'referential consistency'. Unlike traditional criteria of meaning, metalogical referential consistency is not an externally imposed convention, a normative stipulation, an arbitrarily endorsed special rule or standard. Instead, it provides a meta-level criterion or rule of evaluation whose application is *intrinsic*—that is, whose application is realized within the bounds of the framework relativity of individual expressions, statements, and concepts, in relation to their presupposed reference frames. It is a criterion formulated from the standpoint of the metatheory developed here in terms of which the structural and systemic presuppositions of identification of any particular reference frame can be analyzed. A special set of evaluative rules or criteria is not applied across the board in an external way, but, as we have seen, attention is given to those preconditions of identification that must be satisfied in a given reference frame in order for identifying references in that framework to be possible at all. The results of applying such a metalogical criterion of referential consistency are strongly non-arbitrary, both because a special criterion of meaning is not imposed externally, and because such results strongly compel assent—one cannot reject them in a given context of reference without incurring metalogical self-referential inconsistency, which, as we shall see, is equivalent to meaninglessness.

#### 11.4 Formalized description

I am afraid this book will make hard reading.... This is partly due to the strangeness of the subject and the fact that some old ways of reasoning cannot be applied to it.

– Halldén (1949, p. 3)

---

<sup>155</sup> I am not implicitly recommending the formalist's thesis regarding the conventional nature of logical rules. However, the fact that it is possible to view such rules as agreed-upon conventions serves to highlight the contrast between them and the strongly non-arbitrary and compelling criterion of meaning proposed here.

In the interests of clarity, it can be helpful to express referential consistency as a criterion of meaning in more formalized terms.<sup>156</sup> I will not propose here to construct a formal system capable of representing the meta-level principles that we wish to examine; in this book's Supplement, we shall look at the question to what extent this may be possible. Certainly, the development of a formal system is not in itself an indispensable path to conceptual precision. The formalized principles we shall describe are intended to serve as heuristic tools of analysis.

Because some of the notation used in the following is likely to be unfamiliar to readers, the following abbreviated summary of unusual symbols may be helpful as a quick overview in advance; more detailed explanations of each symbol are given later in the text when they are introduced. Familiarity with conventional logical and set-theoretic notation is assumed.

<i>Symbol</i>	<i>Use</i>
$R\alpha o\sigma$	identifying reference established in the form of a ternary relation among a person $\alpha$ , an object $o$ , and a space-time coordinate set $\sigma$
$\vec{p}$	designates a "projective" proposition: vector notation consisting of an arrow over a proposition $p$ indicates that $p$ is metalogically self-referential and that $p$ denies one or more conditions that must be satisfied in order for $p$ to have possible meaning, or for it to be possible to assert, or otherwise use, $p$ meaningfully
$\mathcal{P}$	the set of all putatively true or false sentences (alternatively, propositions)
$\mathfrak{M}$	the set of all varieties of meaning
$\overline{\mathfrak{M}}$	the set of all varieties of meaninglessness
$\mathcal{M}$	a descriptive operator indicating that the sentence or proposition following it is meaningful
$\overline{\mathcal{M}}$	a descriptive operator indicating that the sentence or proposition following it is not meaningful
$\vec{\mathfrak{M}}$	the set of all projective forms of meaninglessness

<sup>156</sup> Readers not at home with formalization should be able to skip to the next section of this chapter without sacrificing continuity.

$\vec{\mathcal{P}}$	the set of all projective sentences (alternatively, propositions)
$M$	designates a precondition of reference that must be satisfied in order for a proposition to be meaningful
$\Vdash$	metatheory derivability sign: metalogical entailment
$\dashv\vdash$	metatheory rejection operator

**Table 11.1 Less familiar logical symbols**

For the sake of simplicity—and without restricting the more general intended scope of inclusion in which non-linguistic varieties of meaning form a part—I limit my treatment here to the set of *putatively referring sentences* (or, alternatively in what follows, *propositions*)  $\vec{\mathcal{P}} = \{p_1, p_2, \dots, p_n\}$  where a  $p_i$  may refer to any one or more  $o_i$  of a set of objects of reference  $\mathcal{O} = \{o_1, o_2, \dots, o_n\}$ , at any space-time location  $\sigma_i$  that is specified by space-time coordinates from the set of coordinates  $\mathcal{S} = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ , and where such a  $p_i$  is *presumed* to possess any significant—i.e., meaningful—truth-value of a set of possible values  $\mathcal{V} = \{0, 1, \dots, n\}$ , where  $n \geq 3$ .<sup>157</sup> By the ‘significant (or meaningful) range of  $\mathcal{V}$ ’ is meant ‘ $\{0, 1, \dots, n-1\}$ ’. (A discussion of the value  $v_n$  follows below.) It is clear that the significant range of  $\mathcal{V}$  is bivalent when  $n-1 = 1$ , with ‘0’ and ‘1’ representing the values “False” and “True”, respectively.

We should explicitly take note in the preceding paragraph of the role of the qualifying phrases ‘putatively referring’ and ‘is presumed’. These qualifications will prove to be necessary given that many allegedly referring sentences are, as we shall see, *mistakenly* believed to be meaningful and either true or false. We shall look at this so-called ‘problem of putative meaning’ in the next chapter.

Some definitions are called for.

D 1            *A particular* is a possible object of identifying reference.

<sup>157</sup> The convention is followed whereby False = 0, and the designated truth-value is  $n-1$ ; the value  $n$  is reserved for a purpose described later.

In the interests of generality,  $p_i$ s with *variable* truth-value may be included: e.g.,  $p_i$ s for which value assignments are a function of time, as may be the case, for example, with future contingent statements.

Alternatively,

- D 2        An *identifying reference* is such that an ascription to that which can be the object of an ascription (namely, a particular) establishes that what is ascribed (one or more properties, relations, a description, etc.) and that that to which the ascription is made are one and the same (identification).

*Comments:*

It is important to recall from {10} that identification may be highly specific, unique to the particular object to which there is identifying reference, or it may be subject to ambiguity, indefiniteness, or probability.

An alternative description of identifying reference was given in {8.6} according to which identification/identifying reference obtains when that which is described is determined as that to which there is reference. This suggests that identification may be conceived as involving both a descriptive component of specification and a component of satisfaction that is present when what is descriptively specified coincides with that to which there is reference. Expressing matters in this way may help to stress the fundamental fact that, when an object is identified, its identifiability is a function of a certain set of criteria that specify parameters within which that object's identification *can* occur, which, in turn, leads inescapably back to one or more reference frames that provide the basis for the possibility of such identification.

D 1 leaves the concept of identifying reference undefined, while D 2 leaves undefined the concepts of particular, description, property, relation, identification, and ascription.

In the interests of economy we will retain D 1, permitting the concept of identifying reference to be primitive from the point of view of formalization. However, it is useful to introduce an interpretation concerning the use of the phrase 'identifying reference':

In what follows, '*R*' is used to stand for an identifying reference, understood as a ternary relation between a person,<sup>158</sup> whose proper name may assume the value of a variable ' $\alpha$ ' ranging over a set of proper names for

---

<sup>158</sup> The inclusion of a person in this relation is provisional here, for formal simplicity and to represent the familiar notion of identifying reference, and is later (see {24–25}) not presumed in every instance of identifying reference.

persons ( $m, n, \dots$ ), a particular (an object of identifying reference)  $o_i$  ranging over a set of objects ( $x, y, \dots$ ), and a space-time coordinate set  $\sigma_i$  ranging over a set of space-time locations ( $s, t, \dots$ ). ' $R_{f_i}$ ' may be read as "reference obtains to ... relative to reference frame  $f_i$ " or "there is reference to ... relative to reference frame  $f_i$ ." It will be desirable to avoid reading ' $R$ ' as "reference is *made* to ..." since such a reading is likely to convey the notion that reference involves an active, agent-based process of referring. As noted earlier in this book, that notion is part-and-parcel of the everyday psychology of referential thinking and speaking according to which references are a kind of causal product of "volitional linguistic acts" on the part of referring speakers. Such a notion brings along with it a menagerie of unanalyzed referential presumptions that comprise, as we shall see later, metalogically objectionable "baggage."<sup>159</sup>

Whenever ' $R$ ' is used, relativity to an adequate framework of reference is understood, from the standpoint of which  $R$  is possible, and so the subscripted ' $f_i$ ' in ' $R_{f_i}$ ' explicitly associates ' $R$ ' with an appropriate reference frame.

We should note that ' $R_{f_i}$ ' is explicitly a meta-level expression: The ' $R$ ' functions as a *meta-frame indicator*. To make this clear, when an individual coordinate point (7,5,-2) is specified, we do not find it necessary to add a layer of psychological-pragmatical interpretation to the effect that a certain linguistic agent has specified that point; it is enough simply to give the coordinates, and in a manner that makes clear what the nature of the coordinate system is in terms of which that point is identifiable.

In the previous paragraph there is identifying reference to the point (7,5,-2). To make that fact explicit, we therefore say  $R(7,5,-2)$ , which informs us that we are engaged in reflection—that we have, as it were, stepped back to consider the specified set of coordinates, and in this sense consider the identified point from a meta-level.<sup>160</sup> In what follows, ' $R$ ' is used in this sense as an indicator that a meta-level standpoint is involved.

When identifying reference (hereafter often simply called 'reference') to a particular obtains,  $R_{f_i} m o_i \sigma_i$  specifies a ternary relation, relative to reference frame  $f_i$ , among the particular  $o_i$  in relation to an individual person  $m$  at space-time coordinate set  $\sigma_i$ :

---

<sup>159</sup> See comments and analyses concerning this extraneous "baggage" in {3.4, 5.4, 23.7, 24.9, 25}.

<sup>160</sup> This is similar, though not identical, to the familiar use-mention distinction. In the text, on the one hand, a certain coordinate point is in view, and, on the other, the subject of attention is the information that the point's coordinates are specified. This latter distinction is more abstract and general, and does not necessarily involve reference simply to the symbols themselves that designate the set of coordinates.

$$(1) \quad (x)(Rmxs \wedge .x \in \{o_1, o_2, \dots, o_n\}: \supset \\ \sim(\exists y)(Rmys \wedge .y \in \{o_1, o_2, \dots, o_n\}: \wedge x \neq y)).^{161}$$

From this point of view, the concept of reference is used to address the metalogical properties of identification—that is to say, possession of an identity is metalogically presupposed in connection with any particular, and by definition all particulars are possible objects of reference, i.e., can be identified.<sup>162</sup>

Let  $p_i \supset R_{f_i}\alpha o_i \sigma_i$  express the claim that a referring sentence  $p_i$  with a value in the significant range implies reference by a person  $\alpha$  at a space-time location  $\sigma_i$  to an  $o_i$ ; in other words,  $R_{f_i}\alpha o_i \sigma$  follows from  $p_i$  whether the value of  $p_i$  is T or F. The claim implicit here is that referring sentences of  $\mathcal{P}$  are such that reference obtains to some  $o_i$  provided only that the  $p_i$ s of  $\mathcal{P}$  have truth-values in the significant range: hence, even when a  $p_i = F$ , reference is considered to obtain to some  $o_i$  that can serve to justify the claim to the effect that  $p_i = F$ .

A  $p_i$  is said to be *self-referentially inconsistent* in three cases which we distinguish here:

(i) When  $p_i \supset R_{f_i}\alpha o_i \sigma_i$  and  $o_i = p_i$ , then  $p_i$  exhibits *sentential* or *propositional self-reference*, depending, respectively, on whether  $p_i$  is considered to be a sentence or to express a proposition. If  $p_i$  is self-referential in either of these two ways and  $p_i$  claims of itself that it is false, then, when  $\mathcal{V}$  is bivalent,  $p_i$  is true iff it is false. Such a  $p_i$  comprises a *paradox-generating* self-referential inconsistency. Many of the semantical paradoxes are clearly of this form.

(ii) When  $p_i \supset R_{f_i}\alpha o_i \sigma_i$  and  $o_i = P_{p_i}$ , where ‘ $P$ ’ designates a pragmatical (or performatory) aspect of the use made of  $p_i$  by  $\alpha$  at space-time location  $\sigma_i$ , then  $p_i$  is termed *pragmatically* (or *performatively*) *self-referential*. If  $p_i$  is pragmatically self-referential and  $p_i$  is such that if  $p_i$  is asserted or otherwise is

---

<sup>161</sup> It follows from this preliminary simplified formulation that identifying reference can obtain between a person and only one object, of a set of possible objects of reference, at a time. The object referred to may, however, be single or it may be compound, as when reference is made to a set having more than one member, or to a set of sets of objects, etc.

From the perspective presented here, when reference to a single object  $o_i$  is determined,  $o_i$  is identified in the sense of (1) in the text. The *identity* of  $o_i$  is essentially a function of  $o_i$ 's identifiability—hence of the framework(s) relative to which reference to  $o_i$  can obtain.

A good deal must be omitted in these preliminaries: The possibility of re-identification would, for example, as noted in {10.1}, need also to be assured.

<sup>162</sup> This is once again reminiscent of Quine's dictum, “no entity without identity,” which we've had occasion to mention a number of times before.

used in a manner so that  $p_i$  is falsified in  $P_{p_i}$ , then, when  $\mathcal{V}$  is bivalent,  $p_i$  is said to be *self-refuting*. The assertion, for example, “This assertion that I make does not refer to an  $x$  such that  $Fx$ ”, for specific values of ‘ $x$ ’ and ‘ $F$ ’, expresses a pragmatically self-refuting self-referential inconsistency. Ramsey’s familiar example, “I can’t say ‘cake’,” when uttered by anyone, accordingly may be seen to be self-refuting.

(iii) When  $p_i \supset R_{f_i}\alpha o_i\sigma_i$  and  $R_{f_i}\alpha o_i\sigma_i \supset R_{f_i}\alpha M_{p_i}\sigma_i$ ,<sup>163</sup> where ‘ $M_{p_i}$ ’ designates a “metalogical precondition of reference” which must be satisfied in order for it to be possible for  $p_i$  to have a value in the significant range, then  $p_i$  is termed *metalogically self-referential*. If  $p_i$  is metalogically self-referential and  $p_i$  is such that  $p_i$  denies one or more conditions that must be satisfied in order for it to be possible to assert, or otherwise use,  $p_i$  meaningfully, then  $p_i$  is said to be *projective*, or to constitute a *projection*,  $\vec{p}_i$ .<sup>164</sup> (In some previous publications,<sup>165</sup> I used the phrase ‘*projective misconstruction*’; in the present study I simply use the term ‘projection’.) Note that we use the vector symbol capped with an arrow to represent projections; the rationale for this choice is explained in the later chapter {13} dealing in detail with projections.

The expression ‘metalogical precondition of reference’ is associated with the following equivalent senses: ‘ $M_{p_i}$ ’ designates a “metalogical precondition of reference” if, in order for reference to be possible in a particular context of reference, the following equivalent conditions hold:  $M_{p_i}$  must be satisfied;  $M_{p_i}$  is a necessary condition of possible reference;  $M_{p_i}$  qualifies as a “metalogical precondition of reference” iff it designates a condition the non-satisfaction of which in a particular context of reference results in projection.

Earlier in this chapter, 18 varieties of meaning were listed in Table B. Let the set of all such varieties of meaning be represented by the German black

<sup>163</sup> I.e., reference obtains from  $\alpha$  at  $\sigma_i$  to the (compound) object of reference  $\{o_i, Mp_i\}$ .

<sup>164</sup> The expression ‘metalogical self-referential inconsistency’ need not be restricted here to the case in which reference obtains to  $\{o_i, Mp_i\}$  at a single space-time  $\sigma_i$ . If  $R_{f_i}\alpha o_i\sigma_i$ , and  $R_{f_i}\alpha M_{p_i}\sigma_j'$ , with  $\sigma_j'$  later than  $\sigma_i$ , and  $\vec{p}_i$ , then we may interpret this as the case in which  $\alpha$  realizes in retrospect that a  $p_i$  endorsed by him or her is projective, i.e., that in endorsing  $p_i$  at  $\sigma_i$  he or she was metalogically self-referentially inconsistent. Analogously, we may have the case where  $R_{f_i}\alpha o_i\sigma_i$ ,  $R_{f_i}\beta M_{p_i}\sigma_j'$ , with  $\sigma_j'$  later than  $\sigma_i$ —i.e., one individual’s commitments can be the basis of another person’s metalogical analysis.

It is important to make a similar distinction in connection with pragmatism self-referential inconsistencies. Statements are sometimes and even frequently made by some individuals who are not aware at the time, and may never become aware, of the pragmatism self-referential inconsistencies which those statements involve.

<sup>165</sup> E.g., Bartlett (1971; 1975; 1976; 1982; 1983; 2005, Part II; 2011, Chapters 2, 8).



letter ‘ $\mathfrak{M}$ ’ and the set of all varieties of meaningfulness by ‘ $\overline{\mathfrak{M}}$ ’. We recall that the set of *putatively* true or false sentences is represented by  $\mathcal{P}$ , and so we should be led to expect a putatively true or false  $p \in \mathcal{P}$  will likely also satisfy  $p \in \mathfrak{M}$ . Let the set of all projective forms of meaningfulness be expressed by ‘ $\vec{\mathfrak{M}}$ ’, so that  $\vec{\mathfrak{M}} \in \overline{\mathfrak{M}}$ , and the set of all projective sentences (or propositions) be represented by ‘ $\vec{\mathcal{P}}$ ’. It follows that  $\vec{\mathcal{P}} \in \overline{\mathfrak{M}}$ .<sup>166</sup>

A metalogically self-referentially inconsistent  $p_i$  makes, with a *putative* value T or F, an ascription A of some object of reference  $o_i$ . Were  $p_i = T$ , then A would apply to  $o_i$ , or  $A(o_i)$ ; were  $p_i = F$ , then  $\sim A(o_i)$ . In either case, possible reference to  $o_i$  would be presupposed:

$$(2) \quad A(o_i) \vee \sim A(o_i) \supset \diamond R_{f_i} \alpha o_i \sigma_i$$

In short,

$$(3) \quad \vec{p}_i \equiv A(o_i) \vee \sim A(o_i) \cdot \wedge \sim \diamond R_{f_i} \alpha o_i \sigma_i,$$

where  $\sim \diamond R_{f_i} \alpha o_i \sigma_i$  is implied by the projective denial of one or more of the conditions that must be satisfied in order for it to be possible meaningfully to assert  $p_i$ .

Consider the case, then, when  $p_i$  is asserted to be true or false, so that  $p_i$  implies that, from the standpoint of reference frame  $f_i$ , reference is established among  $\alpha$ ,  $o_i$ , and  $\sigma_i$ . Hence the possibility of such reference must be assured by relevant metalogical preconditions of reference, and yet it so happens that  $\vec{p}_i$  constitutes a denial of those preconditions. This denial metalogically entails<sup>167</sup> the impossibility of such reference:

$$(4) \quad \vdash p_{i(T \vee F)}, p_i \supset R_{f_i} \alpha o_i \sigma_i, R_{f_i} \alpha o_i \sigma_i \supset \diamond R_{f_i} \alpha o_i \sigma_i, \diamond R_{f_i} \alpha o_i \sigma_i \supset M_{p_i}, \\ p_i \supset \sim M_{p_i} \Vdash \sim \diamond R_{f_i} \alpha o_i \sigma_i$$

Note that the double-bar derivability sign ‘ $\Vdash$ ’ is used to express metalogical entailment. In both ‘ $\Vdash$ ’ and in the subsequently introduced symbol ‘ $\dashv\vdash$ ’ the

---

<sup>166</sup> Implicitly left open by this paragraph is the speculative possibility, not discussed in this work, that for each variety of meaningfulness listed in Table B there may exist a corresponding variety of projective meaningfulness.

<sup>167</sup> For a discussion of metalogical entailment, see the next section in this chapter.

vertical double bar indicates that the operation involved is metalogical, i.e., is an operation belonging to metatheory.

The metalogical self-referential inconsistency of a projection is rendered explicit when the consequent of (2) and the conclusion of (4) are conjoined.

To take an example: P. W. Bridgman's hyperbolic hypothesis to the effect that the entire physical universe is "shrinking homogeneously," i.e., in a manner such that all operations of measurement are correspondingly affected,<sup>168</sup> may be seen to be projective. For the hypothesis to be meaningful, it must in principle be possible for the hypothesis to be known to be true or false. In order for reference to be made to "universal homogeneous shrinkage," Bridgman recognized that the possibility must be presupposed that changes in the relative size of the physical universe could, at least in principle, be detected. This is essential to the meaning of the concept of "cosmic shrinkage." However, by hypothesis "universal homogeneous shrinkage" rules out that the precondition of reference, *possible* detection of the alleged change in relative size, and hence its *possible* meaning, can be satisfied. The hypothesis is therefore projective.

In an intuitive sense, the result  $p_i \Vdash \bar{p}_i$  comes about when  $p_i$  conflicts self-referentially with preconditions that must be granted in order for the value of  $p_i$  possibly to fall in the significant range. A projective assertion consequently involves a special kind of self-referential inconsistency which I have called 'metalogical'. Our main interest here of course is in projective forms of reference.

We say that  $p_i = T$  when

' $p_i$ ' is true iff  $p_i$ (Tarski's definition),

and  $p_i = F$  when

' $p_i$ ' is false iff  $\sim p_i$ .

When  $p_i$  is projective,  $p_i$  is said to have value  $\mu$

' $p_i$ ' has value  $\mu$  iff  $\bar{p}_i$ .

---

<sup>168</sup> Bridgman applied his operational criterion of meaning to the notion of universal homogeneous cosmic shrinkage, originally posed by C. K. Clifford, and concluded that it can have no possible meaning. Cf. Bridgman (1936, pp. 11-12).

Here, ‘ $\mu$ ’ represents the value “projective meaninglessness” which lies outside the significant range of values  $\{0, 1, \dots, n - 1\}$ . It should be clear from the nature of a projective assertion that its value cannot be identified with any of the values in its significant range since one or more conditions are denied which must be satisfied in order for  $p_i$  to have *any* possible value in the significant range. The self-referential inconsistency of a projective assertion is of a kind which literally and logically precludes that the assertion *can possibly* possess a value in the significant range. A projective assertion cannot *possibly* be meaningful. In some contexts there may be a latitude of choice whether to consider an assertion to be meaningless or false (e.g., in the case of the infamous and tiresome “The present king of France is bald”), but from the standpoint of the metalogic of reference, however, no other option is available: The value of a projective assertion *must* fall outside the significant range, hence its value  $\mu$  is equated with meaninglessness.

We shall find that projective assertions—statements that are made which initially are believed to be meaningful, and which are therefore, prior to analysis, believed to be either true or false—are such that simply negating them will also result in projection: i.e.,  $p_i \Vdash \bar{p}_i$  and  $\sim p_i \Vdash \bar{p}_i$ . To take such cases into account we shall need two different kinds of negation: traditional negation, for which I’ve used the symbol ‘ $\sim$ ’ to mean “it is not the case that,” “it is not true that,” or simply “not,” and what for our purposes here will be called ‘*metalogical negation*’, for which—in parallel with our use of ‘ $\mathfrak{M}$ ’ and ‘ $\overline{\mathfrak{M}}$ ’ to designate the set of all varieties of meaning and the set of all varieties of meaninglessness, respectively—we shall employ the script operator  $\overline{\mathfrak{N}}$ : ‘ $\overline{\mathfrak{N}} p_i$ ’ is read ‘ $p_i$  is not meaningful’ or ‘ $p_i$  is meaningless’. Analogously, ‘ $\mathfrak{N} p_i$ ’ is read ‘ $p_i$  is meaningful’.

We note that both ‘ $\mathfrak{N}$ ’ and ‘ $\overline{\mathfrak{N}}$ ’ are descriptive operators in the sense that they convey metatheoretic information regarding the meaningfulness, or the meaninglessness, of the sentence or proposition following it.

However, simply describing and asserting that a sentence or proposition is meaningless does not of itself eliminate it from the domain of coherent rational thought and discourse. The elimination of meaningless statements of the projective variety is one of our main concerns in this study. We shall therefore use the symbol ‘ $\dashv$ ’ to express the *metalogical rejection* of a projective sentence or proposition because it denies one or more preconditions that must be granted in order for it possibly to refer.<sup>169</sup> We’ll read ‘ $\dashv p$ ’ to mean ‘ $p$

<sup>169</sup> Creating this symbol harkens back to Łukasiewicz’s use of ‘ $\dashv$ ’ to mean logical rejection (Łukasiewicz, 1957/1951).

is rationally rejected’, or—since we shall presume rationality<sup>170</sup>—we’ll read ‘ $\neg\parallel p$ ’ simply as ‘ $p$  is rejected’. The rejection operator ‘ $\neg\parallel$ ’ makes explicit that projectively meaningless statements, because they “short-circuit” on the level of possible meaningfulness, are rationally unacceptable, and are rejected for that reason.

In the case above, where  $p_i \Vdash \bar{p}_i$ , we conclude  $\overline{\mathcal{N}} p_i$ , and  $\neg\parallel p_i$ : As a projection,  $p_i$  is not possibly meaningful, and  $p_i$  is rejected. When  $\sim p_i \Vdash \bar{p}_i$ , we similarly conclude  $\overline{\mathcal{N}} \sim p_i$ , and  $\neg\parallel \sim p_i$ .

To make these metalogical operators clear in matrix form:

$\mathcal{N}$	
T	T
T	F
F	$\mu$

$\overline{\mathcal{N}}$	
F	T
F	F
<b>T</b>	$\mu$

The bold ‘**T**’ in the third row below ‘ $\overline{\mathcal{N}}$ ’ tells us that to say of a meaningless statement that it is meaningless is to make a true (and therefore meaningful) statement. We shall return to this observation in a moment.

$\neg\parallel$	
no	T (i.e., the statement or proposition following ‘ $\neg\parallel$ ’ is a member of $\mathfrak{A}$ )
yes	F (i.e., the statement or proposition following ‘ $\neg\parallel$ ’ is not a member of a member of $\mathfrak{A}$ —i.e., it is a member of $\overline{\mathfrak{A}}$ )

The “no” and “yes” in the first column tell us under what conditions  $\neg\parallel$  applies.

---

<sup>170</sup> We shall look more closely at this presumption in {17}.

A  $p_i$  is

(a) *metalogically self-validating* (hereafter simply called ‘self-validating’) in the case in which the *denial* of  $p_i$  (i.e.,  $\sim p_i$ ) is projective ( $\bar{p}_i$ )—i.e., is metalogically self-referentially inconsistent.

Conversely, a  $p_i$  is

(b) *projective* in the case in which the *metalogical rejection* of  $p_i$  (i.e.,  $\neg\|p_i$ ) is self-validating. In short,

$$(5) \quad (x)(x \in \mathcal{P} \cdot \wedge Fx : \supset \cdot G \sim x) \text{ and} \\ (x)(x \in \mathcal{P} \cdot \wedge Gx : \supset \cdot F \neg\|x)$$

where  $F$  is the property ‘... is self-validating’ and  $G$  is the property ‘... is projective’ or ‘... is metalogically self-referentially inconsistent’. A sentence, proposition, or statement that cannot be denied without metalogical self-referential inconsistency then is self-validating, and a sentence, proposition, or statement that is projective is such that its rejection (not its simple negation) is self-validating. The need here for the metalogical rejection operator, rather than simple negation, will be made clear shortly.

Although the problem of putative meaning will continue to hover in the background (until we have solved it in the next chapter)—which has the potential to complicate and confuse examples such as we may contrive to force them to be simple—here nonetheless are two such examples:

- (a)  $p$  : “All identifying references are framework-relative”  
 (is metalogically self-validating)  
 $\sim p \Vdash \sim p$  is projective
- (b)  $q$  : “It is possible to refer beyond the boundaries of possible reference” (is projective)  
 $\neg\|q \Vdash \neg\|q$  is self-validating

We may therefore make the following distinctions:

- metalogically projective *contraries*—as in the case of  $\bar{p}^\rightarrow$  and  $\sim \bar{p}^\rightarrow$ —cannot both be either true or false (i.e., neither can be true or false), but both have the value  $\mu$

- metalogical *contradictories*—as in the case where  $p$  is self-validating and  $\sim p$  is projective—one must be true and the other meaningless (has value  $\mu$ )

It follows that for any  $p_i$  which is such that  $p_i \Vdash \bar{p}_i$  and hence when  $p_i$  has value  $\mu$ , the equivalent claims “the value of  $p_i$  does not fall in the significant range,” “ $p_i$  is not significant,” “ $p_i$  is meaningless” self-validate since the rejection of any one entails metalogical self-referential inconsistency. *For this reason, referential consistency, as a metalogical criterion of meaning, cannot not be accepted.* Metalogical referential consistency is, in other words, a self-validating criterion that must be satisfied in order for claims to be meaningful: It is a necessary condition of meaning.

It may be noted that we’ve left open the significant range of the set  $\mathcal{V}$  of possible values of a  $p_i$ , understanding the significant range to be  $\{0, 1, \dots, n - 1\}$ , where  $n \geq 3$ . Leaving the significant range unrestricted in this way has the advantage of flexibility, since, in some contexts of reference, we may wish to be able to assign such significant values, e.g., representing indeterminacy, statistical probabilities, etc., to a  $p_i$  (for example, in quantum logics). Although no decision has been made, then, in favor of bivalence in  $\mathcal{V}$ , the following metametalanguage formulation is implied by the principle of bivalence, without implying it:

- (i) Every putatively referring sentence of  $\mathcal{P}$  either has a value in the significant range, or it does not.

Adoption of this metalogical version of the principle of bivalence<sup>171</sup> entails that all metalogical statements assigning values from  $\{0, 1, \dots, n - 1, n\}$  — namely, the range of possible values from falsity (0) to one or more designated values ((1) in a classical bivalent system) to  $\mu$ —to a  $p_i$  are themselves either true or they are not. In fact, (i) entails

- (ii) There exist in principle possible procedures which yield a yes or no determination for any metalogical value-assigning statement about members of  $\mathcal{P}$ .

---

<sup>171</sup> For a discussion of the principle of bivalence, a variety of interpretations of the law of the excluded middle, and a related metametalanguage formulation, see Rescher (1969, pp. 148ff).

It will be evident to the reader that the assertion of (i) conjoined with the denial of (ii) constitutes a projective assertion. Consequently, we shall regard (ii) as entailed, in a self-validating manner, by (i).

By way of illustration, let us assume that  $\mathcal{V}$  is three-valued: its significant range then comprises values T (1) and F (0) with  $\mu$  representing the value of projective assertions. The set of sentences or propositions in view is  $\mathcal{P}' = \{p_1, p_2, \dots, p_n\}$ .  $\mathcal{P}'$  includes  $\mathcal{P}$  as a subset;  $\mathcal{P}'$  contains in addition to  $p_i$ s that fall in the significant range,  $p_i$ s that have the value  $\mu$ . For the purposes of assessing referential consistency,  $p_i$  is then three-valued within a bivalent metalanguage.

Matrices for conjunction and negation suitably take the form proposed by Bochvar:<sup>172</sup>

$\sim$	
F	T
T	F
$\mu$	$\mu$

$\&$	T	F	$\mu$
T	T	F	$\mu$
F	F	F	$\mu$
$\mu$	$\mu$	$\mu$	$\mu$

Where  $\mu$  is the value of a projective assertion, the above matrices make clear that the negation of a projection remains meaningless, while the conjunction of a projection with a significant assertion “infects,” so to speak, the compound statement with meaningfulness. The projective character of one

---

<sup>172</sup> A three-valued logic, in which the third value is ‘meaninglessness’ or ‘undefined’, is used by Bochvar to stand for the value of paradox-generating propositions. Although his three-valued system is without a theory of types, it is nevertheless consistent. See D. A. Bochvar (1939, 1943) as well as Church (1939-1940).

Patrick Suppes makes use of Bochvar’s three-valued system (without, however, crediting Bochvar for his truth-matrices) in connection with a formal representation of operationally meaningless statements. Cf. Suppes, (1969/1950).

Several other authors have proposed three-valued systems in which the third value is “meaninglessness”; we shall look at certain aspects of their systems in the next chapter, for example: Halldén (1949), Moh (1954), Åquist (1962). For various reasons, however, special properties of these proposed systems make them unsuitable in the present context.

It might be mentioned that some authors have felt that the matrix for negation ( $\sim$ ) given in the text precludes a satisfactory interpretation of three-valued logic. That  $A$  and  $\sim A$  have the same value when  $A$  has the value ‘meaninglessness’ seems to them problematic. Andrzej Mostowski, for example, has remarked in this connection that he does not have “any hope that it will ever be possible to find a reasonable interpretation of the three-valued logic of Łukasiewicz [which has the same matrix for negation as in Bochvar’s system] in terms of ordinary language” (Mostowski, 1950, p. 223).

It is, of course, my contention that Mostowski’s pessimism was mistaken.

conjunct may undermine the referential consistency of the other conjunct. The matrix for conjunction avoids this potential.

A few pages back the claim was made that we need both simple negation and the metalogical rejection operator in order to deal with the relationship between self-validation and projections; that claim can now be explained: We noted that (i) a sentence or proposition which cannot be *denied* (i.e., negated) without projection is self-validating, while (ii) a sentence or proposition that is projective is such that its *rejection* (not its simple negation) is self-validating. With the above matrix for negation ( $\sim$ ) in view, we see that when  $p$  is projective (its value is  $\mu$ ), merely negating  $p$  (i.e.,  $\sim p$ ) again leads to a meaningless sentence or proposition with value  $\mu$ : negating a projectively meaningless sentence does not lead to a meaningful sentence. Were we therefore to apply negation rather than rejection in (ii), self-validations would have value  $\mu$ . To avoid this unwanted and meaningless result, the rejection operator ( $\dashv$ ) was needed. Instead of negating (or denying) a projective statement, we reject it because it undermines the preconditions of its own possible meaning; we then obtain a statement that is self-validating. We then have a statement that itself cannot be denied (i.e., negated) without metalogical self-referential inconsistency.

Other common connectives are easily defined:

$$A \vee B \quad \text{for} \quad \sim(\sim A \wedge \sim B)$$

$$A \supset B \quad \text{for} \quad \sim(A \wedge \sim B)$$

$$A \equiv B \quad \text{for} \quad (A \supset B) \wedge (B \supset A),$$

so that the following matrices are determined:

$\vee$	T	F	$\mu$
T	T	T	$\mu$
F	T	F	$\mu$
$\mu$	$\mu$	$\mu$	$\mu$

$\supset$	T	F	$\mu$
T	T	F	$\mu$
F	T	T	$\mu$
$\mu$	$\mu$	$\mu$	$\mu$

$\equiv$	T	F	$\mu$
T	T	F	$\mu$
F	F	T	$\mu$
$\mu$	$\mu$	$\mu$	$\mu$

From these matrices it can readily be seen that once part of an expression assumes the value  $\mu$ , the expression automatically assumes the value  $\mu$ . (The same rationale applies here as in the case of conjunction.)



It is also evident that if all  $\mu$ -rows and  $\mu$ -columns are *eliminated*, the matrix is reduced to the normal two-valued one. If one sets  $T, F = \Psi$ , then it is clear that the elimination of statements of value  $\mu$  leaves a set of statements having the value  $\Psi$ , statements which are in the significant range. This is an obviously desirable property of a necessary, not sufficient, criterion of meaning: Its application will lead to the elimination of certain meaningless statements, leaving untouched all candidates which may be significant (and perhaps to which additional necessary conditions of meaning apply.)

The preceding discussion leads us to affirm the following metalogical heuristic principles<sup>173</sup> that will play a central role in later analyses:

- |     |   |  |
|-----|---|--|
| (6) | $\vec{p} \Vdash \neg p$   | If $p$ is projective, this metalogically entails that $p$ is rejected  |
| (7) | $\neg p \Vdash p \notin \mathcal{P}$                                      | The metalogical rejection of $p$ entails that $p$ cannot (even putatively) be considered a true or false sentence or proposition |
| (8) | $\vec{p} \Vdash p \in \overline{\mathcal{M}} \vdash p \notin \mathcal{M}$ | If $p$ is projective, it is meaningless, i.e., $p$ is not included in the set of meaningful sentences or propositions            |
| (9) | $\sim[\vec{p} \Vdash \sim p]$   | The fact that a sentence or proposition is projective does not metalogically entail the truth of its negation                    |

---

<sup>173</sup> As noted at the beginning of this section, such principles are intended as useful and convenient heuristic tools, as metatheoretical rules to guide subsequent analyses. We have no interest in whether they are logically independent of one another or in developing a formalized deductive system here.

- (10)  $\vec{p} \Vdash \neg\|p \wedge \neg\|\sim p$  If  $p$  is projective, both  $p$  and  $\sim p$  are rejected
- (11)  $\vec{p} \Vdash \overline{\mathfrak{N}} p \wedge \neg\|p$  If  $p$  is projective,  $p$  is meaningless and is rejected
- (12)  $\overline{\mathfrak{N}} p \Vdash \mathfrak{N}(\overline{\mathfrak{N}} p)$  To say that  $p$  is meaningless entails that it is meaningful to say this

The metalogical criterion of meaning that emerges from this discussion is both *strongly non-arbitrary* and *compelling*. It is *strongly non-arbitrary* because the criterion is intrinsically informed by the specific nature of individual contexts of reference, that is, by the specific constitutive structure of individual reference frames. It is *strongly compelling* because one cannot at one and the same time consistently use expressions, sentences, or concepts referringly yet undermine their capacities to refer. Finally, a metalogical criterion of meaning defined in terms of referential consistency is *self-validating*; denying its application leads to projection.

In such a metalogical understanding of meaning, criteria for evaluating referential consistency and meaning are determined as a function of one's needs and interests in referring to the kinds of objects under consideration. Within any specific context of reference, with these needs and interests in view, intrinsically determined criteria for evaluating referential consistency and meaning *merge* from the standpoint of the metalogic of reference. They provide critical tools for appraising the meaningful use of expressions, sentences, or concepts in that context. Referential consistency is, in short, a contextually determined, yet strongly non-arbitrary, compelling, and self-validating criterion of meaning.

### 11.5 Metalogical entailment

When dealing with highly abstract analysis such as we undertake from a metalogical level, there is a temptation to think that the results of such metatheoretical, reflective analysis can be applied retroactively, so to speak, to the original subject-matter as given prior to analysis, a subject-matter that we are now able to view from the standpoint of a higher order, more inclusive, analytically self-conscious frame of reference. In one way, such a retro-

active application makes sense, and in another, it does not. Elsewhere (Bartlett, 1974, 1975a), I distinguished between these two levels in terms of content or information that is *implicit* as opposed to that which becomes *explicit* through reflective analysis. The relationship between implicitly given content or information and that which we come to see by means of meta-level analysis can be complex and misleading. Since much of the discussion in this chapter has presupposed this relationship, some words about it are appropriate.

A central concern in this work, as the reader by now should see clearly, is to make explicit those structural/systemic presuppositions without which any frame of reference cannot possibly function in identifying the set or sets of objects within its scope. Once we have made such presuppositions explicit, we frequently are inclined to give the results of our analysis retroactive application or validity—which is to say that we come to believe that those results were “there already, implicitly,” that they should have been recognized and affirmed by any rational being from the very outset—in other words, that, although we have had to go through sometimes laborious steps of analysis to reach those results, if we have in fact reached conclusions that cannot *not* be accepted without metalogical self-referential inconsistency, then anyone who is rational is compelled to accept those conclusions, since those conclusions were *implicitly entailed* by the subject-matter we began with.

As I have shown in Bartlett (1974, 1975a), that inference is projective: it entails metalogical self-referential inconsistency; the meta-framework in terms of which those results are reached is—and this should now be evident—simply not available until we have situated analysis on a meta-level and gone through an abstract, reflective process of reasoning.

Since this matter is crucial to an understanding of what we seek to do in this work, it is worthwhile to make this clear on a concrete level. Consider much the same issue from the simplified perspective offered by Jaakko Hintikka when he sought to explain his notions of “internal criticism” and “immunity to criticism”:

[S]uppose a man says to you, “I know that *p* but I don’t know whether *q*” and suppose that *p* can be shown to entail logically *q* by means of some argument which he would be willing to accept. Then you can point out to him that what he says he does not know is already *implicit* in what he claims he knows. If your argument is valid, it is irrational for our man to persist in saying that he does not know whether *q* is the case.... (Hintikka, 1962, p. 21, italics added)

Hintikka goes on to speak of the “immunity to criticism” that a demonstration of this kind can claim for itself: Although he doesn’t express it this way, the general value of such a demonstration lies in the fact that certain results have been shown to be *entailed implicitly*<sup>174</sup> by an originally accepted claim, position, or argument. “Immunity to this kind of criticism (or persuasion) seems to me a notion important enough to deserve serious study...” (p. 31).

Hintikka mentions “internal criticism”—a way of showing, by means of rules that have been set down, or by means of other techniques, that “a sentence can be made true by the sole means of internal criticism” (p. 36). He does not develop this notion of “internal criticism” beyond the following perceptive observations:

If  $q$  is entailed by  $p$ , then the state of affairs expressed by  $p$  cannot be realized without realizing the state of affairs expressed by  $q$ , too. But from this it does not follow, obviously, that anybody who knows that  $p$  should for this reason actively know that  $q$ , unless it be assumed that he is making the best possible use of his knowledge. Logical truths are not truths which logic forces on us; they are not necessary truths in the sense of being unavoidable. They are not truths we *must* know, but truths which we *can* know without making use of any factual information. The logical implications of what we know do not come to us without any work on our own part; they are truths which we can extract, often with considerable labor, from whatever information we already have....

The applicability of our results may thus be said to presuppose a certain amount of rationality in the people whose attitudes are being discussed. (pp. 37-38).

Two thoughts stand out from the passages just quoted: that valid, logically compelling reflective analysis may be said to have a relation of what I’ve called ‘implicit entailment’ in relation to the original subject-matter that is the focus of analysis, and that the acceptability and indeed the persuasive force of the results of such analysis presuppose rationality on the part of any audience.

These are important points to bear in mind, but they do not quite serve our purpose, which is to underscore that two fundamentally distinct reference frames are involved in what Hintikka has described: There is what we might call ‘the pre-analytical reference frame’, and then there is what we may term

---

<sup>174</sup> Hintikka (1962, p. 32) calls this “virtual implication,” a misnomer since entailment is really involved (as he apparently recognized in the passages quoted from pp. 21 and 37).

‘the reference frame presupposed by the analysis itself’. The results reached from the standpoint of the second, reflective, framework are not simply transferable back to the pre-analytic standpoint. The constitutive structures of the two frameworks are fundamentally distinct, and are not equivalent. The analytical results reached are a function of the meta-frame in terms of which they are possible; we need to remind ourselves of this framework relativity.

In this work, when I speak of “metalogical entailment,” or simply “entailment” that presupposes a metalogical level of analysis, the notion of “implicit entailment” always resides in the background. We need regularly to remind ourselves that the metalogical results reached presuppose a certain maximally general meta-level of analysis, and cannot in a theoretically naive fashion be applied retroactively.

## 12

### **The Problem of Putative Meaning and the Logic of Meaninglessness**

Underlying the familiar distinction between truth and falsehood, there is a more radical distinction between significance and meaninglessness. True and false statements are both significant, but some forms of words, with the vocabulary and construction of statements, are neither true nor false, but nonsensical—and nonsensical not for reason of wording or of grammar, but for logical reasons.

– Gilbert Ryle (1951, p. 4)

[A]s of today, after thirty or forty years of use of the concept of meaninglessness, we still have no adequate, general criterion for proving that a statement is meaningless.... The fact is that we have no reliable, general test of what is “meaningless” (in the sense of “neither true nor false”).

– Edward Erwin (1970, pp. 25-26, 161)

The general concern to identify and eliminate meaningless concepts and statements from technical and even from ordinary discourse reflects a long tradition in which logic and philosophy together have sought to clarify our conceptual structure and exhibit departures from sense. For example, in a letter to Lambert, dated September 2, 1770,<sup>175</sup> Kant made mention of the need for a “negative science,” a *phaenomenologia generalis*, which would undertake what might today be construed as the construction of a kind of “sieve” capable of sorting out the meaningful from the meaningless, a negative science whose object would be to insure that only meaningful concepts, propositions, and statements remain as the subject for subsequent analysis and

---

<sup>175</sup> Kant (1999, pp. 107-109).

potential use. The list of names of philosophers who have sought, directly or indirectly, to contribute to such a goal could be expanded almost indefinitely.

During the early twentieth century, several groups of philosophers were strongly influenced by the appeal of this approach: The Vienna Circle urged the acceptance of verifiability as a criterion of meaning supportive of the objectives of science. Some years later, Nobel Prize winning physicist and philosopher P. W. Bridgman advocated operationalism as a criterion to weed out the meaningful from the meaningless, while Oxford philosopher Gilbert Ryle directed attention to category mistakes that he believed were at the basis of important philosophical problems. All shared an interest in purging philosophy of forms of alleged and blameworthy meaninglessness.

Their efforts, however, gradually lost headway as criticisms compounded over the years that showed that justification for the proposed criteria of meaning was not compelling in the strong sense discussed in the previous chapter. Moreover, many philosophers refused to accept that the non-satisfaction of the criteria of meaning that had been proposed truly resulted in meaningless statements, propositions, or philosophical views.

The historical failure of the criteria of meaning that had been recommended led during the second half of the twentieth century to a widespread decline of interest on the part of philosophers to seek for the solution to perennial philosophical problems by means of conceptual tools designed to detect meaninglessness. This was not only a decline of interest, but an emotionally understandable *resistance* to the very notion that many philosophical questions and problems, in which professional philosophers had a vested interest, should find their solution in the recognition that forms of meaninglessness were at their core.

As a professional group, many philosophers had already become defensively sensitive to the facts discussed in Part I of this book, relating to the lack of established results in philosophy after two and a half millennia of effort. For them, the suggestion that forms of unacknowledged meaninglessness might be responsible for the discipline's lack of constructive progress could of course be anticipated to arouse ire and antagonism. This was already evident some 50 years ago when Erwin (1970, p. 127) remarked: "In giving ... a systematic account [of the concept of meaninglessness], one must, I think, deal with the skepticism and hostility a growing number of philosophers feel toward the use of this concept in solving philosophic problems." Writing at the same time, Weiler (1970, p. 328) reflected what had by then become the prevailing consensus: "That there are no universally valid criteria of meaning has been demonstrated.... What we give meaning to and what we deny

meaning to in a given situation depends on our interests, or ... on our point of view.”

This relativist view of meaninglessness immediately leads to the blurry notion that there is no “real meaninglessness,” since a meaning—for any symbol or sequence of symbols, for any asserted statement, for any proposition whatever—can always be found, or be constructed, or be imagined, and hence it would seem to follow that there is nothing that is “in and of itself” meaningless. This uncertain situation was imagined even by P. W. Bridgman, the stalwart defender of operationalism:

A ... paradoxical and embarrassing situation arises whenever we make a statement of the form “The statement A has no meaning.” For if the statement A did not have a meaning of sorts, we would not be able to assent to the statement that it has no meaning. We may recognize this “meaning of sorts” as a second kind of meaning, which may be defined in terms of the response elicited when the statement is made. If the response elicited is always the same or if the different responses have recognizable elements in common, then it may be socially useful, and it may be in accord with the usual implications of language, as used, to admit a second sort of meaning. The primary meaning of meaning would then be sought in the purposes and operations leading to the usage, and the second meaning in the response actually elicited in the given cultural context by the particular verbal combination. We might be able to get along with this sort of situation, and I think popular usage as a matter of fact does, but we would have to admit such questions as: “What is the meaning of the meaningless statement A?” Although perhaps possible, it seems to me that it is too confusing to admit this second sort of meaning, and I shall endeavor to find other ways of talking. (Bridgman, 1959, p. 34)

Perhaps with Bridgman we are free to choose “other ways of talking,” but are we, in the process, simply sweeping a problem under the carpet, where, out of sight, we are no longer troubled by it? Is the meaningless always or frequently, in some secondary sense, meaningful?

Later, in the same book, Bridgman reflected:



[I]f my neighbor responds to my communication in the way that I wanted him to respond, it would seem pretty natural for me to say that what I had said had some sort of meaning and my neighbor knew what it was, even though both of us might have been talking about some such thing as absolute length or absolute simultaneity, which from the operational point of view of the physicist is meaningless. In this sense my neighbor and I can mean something by a statement, even if it can be shown that we are both deceiving ourselves. (Bridgman, 1959, p. 213)

If it can be shown that Bridgman and his neighbor are both “deceiving themselves” in taking as meaningful something which is not, then should we not charitably presume that they must have had something in mind which at least seemed at the time to be meaningful to them? This was a view examined by Avrum Stroll (1955), a view he found expressed by fellow philosopher W. W. Mellor. Mellor (1954) proposed that people really have something in mind, despite the fact that what they believe may be meaningless, and what they have in mind is really described by describing a state of their mind.<sup>176</sup>

Today, three-quarters of a century later, it remains a logically perplexing problem how it is possible for anything—whether thought or said or written—to be truly and correctly judged to be without meaning. Even what we call ‘gibberish’ would seem to have some meaning—the meaning, that is, that we associate with “gibberish.” To claim that something is truly meaningless would seem to be a rather extravagant claim: Perhaps when one makes such a claim, one is really, at base, expressing what is simply an exaggeration, or perhaps one is only venting impatience in the face of what is in some way judged undesirable or objectionable?

In an effort to define this issue more strictly, should we perhaps rather consider the possibility that some things that are thought, said, and written are “incoherent” in a way that *precludes* that they *could* possess a meaning. But are we really capable of having such thoughts, saying such things, or writing such things that literally and logically “self-destruct” on the level of *possible* meaning? And, even if there exists such an outlandish self-undermining human capability, isn’t there a sense—a meaning—that ought, if only out of simple charity, to be associated with the *intended* thoughts, the *intended* sayings, the *intended* writings—before they were recognized as incoherent and senseless? Shouldn’t we—not only out of charity, but out of intellectual honesty—grant that such things, which we may later come to realize are so much

---

<sup>176</sup> Stroll rejected this view, for reasons not relevant here.

nonsense, *must have had* at least *some meaning* for those who intended them to have such meaning as they were mistakenly believed to have?

If we take these questions seriously, we face the vaguely defined outlines of the so-called ‘problem of putative meaning’.

### 12.1 The problem of putative meaning

[T]he job of being rigorously rational even about irrationality  
... is just not done.

– Arthur Prior (1961, p. 31)

Meaningfulness which is “putative”—that is, which makes sense only allegedly—is closely associated with seeking to be “rigorously rational” even about what is devoid of rationality. To take seriously the issue posed by putative meaningfulness is perhaps to embrace a measure of intellectual honesty, but the matter can be theoretically elusive. Arthur Prior had in mind the logical paradoxes when he made the statement quoted above. But the issue extends, as logician J. L. Mackie noted, beyond the paradoxes, to include “irrational thinking,” and “states of affairs that seem in themselves to defy reason” (Mackie, 1973, p. 295).

Perhaps, at its most fundamental level, the problem of putative meaning reduces to the question how we are to handle—in a way that makes sense—statements, propositions, and concepts that we determine to be meaningless, while all the while taking into account that such statements, propositions, and concepts are frequently regarded as meaningful by people who do not realize in a strictly rational way what they are thinking and saying. As Australian philosopher Leonard Goddard put the issue: “Nonsense that can be understood is a peculiar sort of nonsense” (Goddard, 1964, p. 313).

It will help us to advance beyond these informal and vague remarks by means of a brief return to formalization.<sup>177</sup> In what follows, I continue to use the notation introduced in the previous chapter.

Let  $p_i$  be a sentence, statement, or proposition in the context of a system SI which permits unambiguous identifying and re-identifying reference<sup>178</sup> to a set  $\mathcal{O} = \{o_1, o_2, \dots, o_n\}$  of objects. Let it further be agreed that a  $p_i$  is asserted to have a truth-value in the significant range, i.e.,  $\neq \mu$ .

<sup>177</sup> This chapter section is based in part on Bartlett (1982), with several emendations.

<sup>178</sup> An explanation of the rationale for the condition requiring that re-identification be possible in SI will be found later in the third from last paragraph of this section.

Upon analysis, it is determined that  $p_i \supset \bar{p}_i$  because  $p_i \supset R\alpha o_i \sigma$ , while  $Mp_i \supset \sim \diamond R\alpha o_i \sigma$ . From a metalogical frame of reference,  $f_i$ , then, we associate with  $p_i$  the truth-value  $\mu$  not in the significant range. Note that *this* claim is an assertion about  $p_i$ -in-SI, and hence is a metalogical claim whose truth-value is determined on a bivalent metalogical basis.

In this context, the problem of putative meaning poses itself in a clear way: Opponents to the use of meaning criteria have argued that, on the one hand, we have an expression, sentence, or concept that is used in various contexts and in what is considered to be a meaningful fashion. Yet, on the other hand, upon application of some particular criterion of meaning, the alleged meaning is supposed to be given up, and the matter closed. The initially perceived meaning is, as the proponent of the given criterion of meaning claims, to be judged “not really meaningful” and dismissed. However, to insist upon this runs counter to the belief of those who regarded the given expression, sentence, or concept to be meaningful; for them, the insistence by the proponent of the criterion of meaning is perceived as literally (not logically) paradoxical.

However, the quasi-paradoxical appearance of the problem of putative meaning is easy for us now to dispel: From the standpoint of SI,  $p_i$  is used to refer to an  $o_i$  so that  $o_i$  is uniquely determined. From the standpoint of  $f_i$ , reference is made to  $p_i$ -in-SI and reveals, through an analysis of  $p_i$ 's referential preconditions, that the assertion of  $p_i$ -in-SI undermines  $p_i$ 's capacity to refer to  $o_i$ .

If we associate a “meaning spectrum”  $V'$  with  $p_i$  such that  $V' = \{0, 1, \dots, n\}$ , where  $n = \mu$ , then for any  $0 \leq v_i < n$ ,  $v_i$  falls within the significant range  $V'$  of  $V'$ . While the assignment of any  $v_i$  up to and including  $v_{n-1}$  may be made from the standpoint of SI,  $\mu$ -assignments require recourse to a metalogical frame of reference  $f_i$ . In short, the possibility of detecting that a  $p_i$  has value  $\mu$  is essentially a function of  $f_i$ 's referential capacity. A metalogical statement  $S$  asserting that  $p_i$  is projective in SI, independently of  $f_i$ , itself is projective, as the reader may confirm.

From this reflective standpoint, the problem of putative meaning does not arise. What opponents to the use of meaning criteria very likely wish for by way of metatheoretical explanation appropriately falls under the heading of “making mistakes” and “detecting errors.” When one makes a mistake without realizing it at the time, and later discovers his or her error, the passage of time provides what is, in effect, a metasystem that permits reference to what is retained in memory: From this vantage point, one compares what one remembers having thought earlier with what one now knows, and claims, in retrospect, that a mistake was made at the earlier time. The same may be said in

the present case: The use of  $p_i$  putatively to refer to  $o_i$  in SI was *erroneous* because  $p_i$  can be shown to be projective in  $f_i$ . Hence, to make an assertion that can be shown to be projective and hence meaningless in the sense developed, is to make one of many different kinds of possible mistakes.<sup>179</sup>

To remind us of this, it is convenient to view  $\mu$ -assignments as involving, in a very literal sense, a *shift of significance*. Assumption of a metalogical frame of reference with respect to a projective assertion  $\bar{p}_i$  results in a shift in  $p_i$ 's *putative* truth-value (in SI) to  $\mu$  (in  $f_i$ ). Such a shift in significance is essentially a function of the metalogical frame of reference used. We shall later find as we explore specific applications of the metalogic of reference that they reveal many such *shifts to the value*  $\mu$  of expressions, sentences, and concepts erroneously believed to be significant.

The problem of putative meaning is resolved in this way, and in a manner that should satisfy opponents to the use of meaning criteria. By recognizing the framework relativity of the metalogical analysis that identifies a given  $p_i$  as projectively meaningless, the pre-analytically accepted "meaningfulness" of  $p_i$  retains whatever sense was, prior to analysis, it was believed to possess, but that "sense" is now placed in scare quotes because we realize in retrospect, from a metalogical point of view, that this "sense" was—and I know no better term for it—"delusional."

## 12.2 The delusion of meaningfulness

As a species, and not only as philosophers, we are prone to a wide assortment of errors of judgment and of reasoning, to sensory and cognitive illusions, and to a capacity to deceive ourselves in a wide range of circumstances. To suggest that we also do at times mistake what is fundamentally meaningless to be meaningful is not a particularly radical observation to make, and yet to propose that many perennial philosophical questions and problems have at their core mistaking what is fundamentally meaningless for what is meaningful is indeed a radical and revolutionary step. Despite anticipated groans, frowns, and headshaking of those of many of my fellow philosophers who have a deeply seated investment in regarding such questions and problems as meaningful, the present work advances the claim that a particular and conceptually resistant variety of *delusion of meaningfulness* has impeded constructive philosophical progress. I use the term 'delusion' not in a metaphorical sense, but

---

<sup>179</sup> Nothing need be said in any detail about what one's "intentions" may have been in using  $p_i$  in this way, since referring to what one had in mind but sees was not realized in actual expression, is in practice to orient oneself with respect to  $p_i$ -in-SI in the manner already described.

rather, as will be made clear, in its conventionally accepted psychiatric meaning.

I have used the terms ‘erroneous’ and ‘mistaken’ when applying them to the belief that one may have in the meaningfulness of something that is later revealed to have no possible meaning because the preconditions for its possible meaningfulness are not satisfied, or those preconditions may in fact be explicitly rejected. To call such beliefs ‘mistaken’ or ‘erroneous’ is certainly appropriate but not, to my mind, as informative as is the more suitable term ‘delusional’.

There is an interesting and instructive relationship that may be drawn here between, on the one hand, the philosophical focus in this section on the problem of putative meaning, and, on the other, an area of study in psychiatry that concerns what have been called ‘*delusions of misidentification*’. I shall charitably save the reader from a detour in the psychiatric theory of such delusions.<sup>180</sup> It will be enough for our purposes here only to sketch the psychiatric definition of such delusions, which also are non-metaphorical and which relate—with an underscored *mutatis mutandis*—to our present discussion of putative meaning.<sup>181</sup>

Characteristic of delusions of misidentification are erroneous and firmly established beliefs which have two general, distinguishable components that apply directly to the problem of putative meaning: a delusional belief coupled

---

<sup>180</sup> A description of delusions of misidentification and a discussion of their connection with certain of the author’s studies in epistemology will be found in Bartlett (2005, Chapter 19, “Pathologies of Everyday Thought,” § Delusions of Misidentification).

<sup>181</sup> Despite how apropos the phrase “delusions of misidentification” might appear to be in the context of the present study, these delusions make up an area of interest to psychiatrists and clinical psychologists that is poles apart from the concerns of the present study; this psychiatric species of delusion has no direct connection either with the problem of putative meaning or with the epistemological dimensions of reference that define the subject-matter of this study.

For the incidental information of the curious reader, psychiatric delusions of misidentification comprise a group of mental disorders in which there is a firmly rooted but mistaken belief in the identity of other people, oneself, places, or objects. Some of these delusions were mentioned in passing in {3.3}, and there are others. They include the Cotard delusion (the patient believes he has died), the Fregoli delusion (he is being followed by people he knows, but they’re in disguise), the mirrored-self misidentification (the patient’s image in a mirror is not his own), reduplicative paramnesia (someone who has died is still present), unilateral neglect (part of the patient’s body belongs to someone else), alien control (his behavior is being controlled by someone else), thought insertion (someone is introducing thoughts into his mind), delusions of reference—again, having no connection with the metalogic of reference—in which the patient believes that the actions of others have some special reference to him), and there are other varieties. For a more complete discussion, see Coltheart & Davies (2000), Joseph (1986), Christodoulou (1991), Roberts (1991), Gold & Hohwy (2000), Sims (2003), and Bartlett (2005, Chapter 19).

with an erroneous identification. The erroneous identification is of course the mistake that is made when what is fundamentally devoid of meaning is nonetheless considered to be meaningful. The delusional belief is, in these circumstances, the conviction of meaningfulness. We shall find that both of these play a clear and central role in our later analysis of projections.

The definition of the term 'delusion' that is conventionally accepted by psychiatrists and by many clinical psychologists has been legislated by the *DSM*, the *Diagnostic and Statistical Manual of Mental Disorders*, the diagnostic Bible of psychiatry. It defines *delusion* as meeting four requirements, which I abbreviate here: A psychiatric delusion involves (i) an erroneous belief which is (ii) rigidly adhered to, (iii) in spite of what others in the prevailing society believe, and (iv) despite incontrovertible evidence to the contrary. When such delusional beliefs become walled off from other beliefs which are rationally based and which are held by the same individual, they develop "*cognitive impenetrability*," as Pylyshyn (1984) aptly named this phenomenon. That is to say, delusional beliefs acquire an *immunity to revision* in the light of conflicting evidence, and despite the possible presence, in the same individual, of other accurate, reality-based beliefs that are in fact based upon, and do acknowledge, such conflicting evidence.

Additional defining properties of delusions in the extended, epistemological sense I shall employ, not incorporated in *DSM's* definition, include the fact that delusions are *exempted from criteria of consistency* with other beliefs that are based on actual evidence; they are *exempted from criteria of rationality*; and they are accompanied by what has come to be known as '*anosognosia*' or *lack of awareness of the existence of delusion*.

Delusions of meaningfulness involve, in addition to the erroneous belief already noted, a second component of erroneous identification, the basis for the mistake that the delusional belief is built upon. As we've seen in previous chapters, identification is necessarily framework-relative; in the case of delusions of meaningfulness, the associated presupposed background framework of reference is frequently that of the general conceptual structure of the individual or group, commonly understood by psychologists and social scientists as the product of both individual and group psychology, and of social and environmental factors. This general conceptual framework tends to serve as the shared context in terms of which people think and interpret events. In the case of philosophers, framework relativity in terms of a general conceptual structure shared by practitioners of the discipline plays a similar role; the reference frames of philosophers and of their distinguishable disciplinary sub-groups are simply more specialized and technical than that of the general society. If delusions of meaningfulness afflict philosophers in connection with certain of

the questions and problems they study, the framework-relative context of such delusions is likely to be found in the common ground of their shared conceptual structure.

Whatever people identify as meaningful is a function of a background frame of reference whose constitutive structure is seldom explicitly subjected to reflective analysis. Because their presupposed conceptual framework is relied upon implicitly and habitually, the relativity of misidentifications to that presupposed reference frame tends to be overlooked, neglected, and, especially important here, it is frequently *transgressed*. It is here, as we shall see, that specifically philosophical/epistemological delusions of meaningfulness find their basis.

In later chapters, we shall explore a number of such transgressions; we shall find that all exhibit the susceptibility of philosophers and of non-philosophers alike to mistake as meaningful what is, upon reflective analysis, self-undermining on the level of *possible meaning*. We shall find that mistakes of this kind qualify as real, non-metaphorical delusions of meaningfulness, to which believers strongly adhere in a way that walls off their delusional beliefs from revision, exempting those beliefs from standards of consistency and rationality. We shall find that here, in such delusions of meaningfulness, lie the main obstacles that have been responsible for philosophy's failure to realize constructive progress during the past two and a half millennia.

### 12.3 The logic of meaninglessness

Logic deals with the truth-relationships among propositions. But before there is truth there must be meaning. And the *bete noir* [sic] in this regard is meaninglessness.

– Nicolas Rescher (2014, p. 63)

Readers will by now be aware that the boundaries dividing meaning from meaninglessness define one of our central interests. If, as this work seeks to justify, a significant number of philosophical questions and problems find their resolution—that is, are dissolved and disappear—through a recognition that they “short-circuit” on the level of possible meaning, the logic of meaninglessness should be expected to play a role of central importance. In this section, we shall look briefly at some of the work in this field that has been done in the past.

Logicians who have published major book-length studies, book chapters, or even individual papers about the logic of meaninglessness are few and far between. This study would be incomplete without some discussion of certain of their works and the place of this study in relation to them. To this end, I do not propose to give a historical commentary of the handful of logical studies in this area, but instead will highlight a number of the most important and potentially relevant features of their formalized accounts.

As might be expected, much of the work on *logics* of meaninglessness—for a number of different systems have been proposed—was closely associated with and received its motivation from proposals by philosophers such as Schlick, Ayer, Carnap, and Reichenbach, who, as we saw in the preceding chapter, formulated and urged the acceptance of their respective criteria of meaning. Their proposals occasioned the interest on the part of a group of logicians to develop formal logics capable of dealing with the class of meaningless statements<sup>182</sup> comprised of those statements that fail to satisfy a given criterion of meaning. Some of these philosophers developed formalized systems, others gave formalized accounts that did not aspire to the status of deductive systems and yet sought in one way or another to accommodate the resulting meaningless statements.

The motivation to do this was clear: All of these researchers urged a more rigorous, analytically self-conscious use of language; it was evident to them that language can be used meaningfully or in ways devoid of meaning; it followed in their views that a comprehensive formal logic should take into account not only that portion of human discourse that is meaningful, but also that part that fails to express meaning. As a result of their language-based focus, their shared tendency was to define meaningfulness in terms of statements that are either true or false: If a statement is true, it is meaningful; if false, it, too, is meaningful; if it is neither true nor false, in the views of these researchers, it is not meaningful. As we shall see, this truth-functional conception of meaning is radically distinct from the conception of meaning with which we are concerned in this study.

The systems of logic and formalized approaches that were developed to take into account meaningless statements were, with a few exceptions, largely three-valued: they included the values “true,” “false,” and “meaningless,” or sometimes called ‘undefined’ or ‘indeterminate’.

Russian logician Dmitri Anotolevich Bochvar (1939, 1943), whose matrices were described in the last chapter, formulated a three-valued logic in which the third value is “meaningless” or “undefined.” He used this third

---

<sup>182</sup> In what follows in this section, unless otherwise indicated, “statements” may be regarded indifferently to include sentences, statements, or propositions.



value to stand for the value of paradox-generating propositions. Other logicians who have worked on logics of meaninglessness, best-known among them, Swedish Sören Halldén (1949), Chinese Moh Shaw-kwei (1954), and Swedish Lennart Åqvist (1962), also proposed three-valued logics in which the third value that statements might take is “meaningless” (or, again, “undefined” or “indeterminate”). They often called their work ‘logics of nonsense’ rather than ‘logics of meaninglessness’, but this was only a choice of words.

Of these logicians, Halldén undertook what is the most comprehensive study of a three-valued logic in which the third value was designed to accommodate meaningless statements; of the three authors mentioned, Halldén is the only one who devoted a full-length book to the subject. We shall describe his approach in more detail in a moment.

Moh, like Bochvar, interpreted the third value in his three-valued logic to represent the value of paradoxical statements. He employed a three-valued system developed by Łukasiewicz, which is not appropriate for our purposes here since in that system the conjunction of a meaningless statement with a false statement is considered to be false.

Instead, it has never made sense to me to consider a compound statement—whether it is, for example, a conjunction, a disjunction, or an implication—to be meaningful if one of its terms is itself devoid of meaning. In the case of a compound statement, it is its meaning *as a whole*, and not the possible meaning of some of its parts, that is significant. In the case of projective meaninglessness, where the very possibility that a statement can possess a meaning is undermined, it makes no sense to form a compound statement in which a projectively meaningless statement forms a part, and then to consider the result to be meaningful. If a statement is logically compound, consisting of two or more statements joined by logical connectives, the compound statement as a whole is meaningful if and only if its component statements are meaningful.

Presley (1961, p. 232) offers an example of those who reason differently about this. “It appears that there is no reason why  $S \vee Q$  should be ruled out as meaningless when  $S$  is meaningful and  $Q$  is not.”<sup>183</sup> But there *is* a reason one might give: The reason that I myself should give is that  $S \vee Q$  is a compound statement, which as such has no meaning as a whole—in this case as a disjunction—when  $Q$  is meaningless. It makes no rational sense to assert a disjunction one of whose disjuncts has no meaning.

But the issue here is somewhat more complex, for to maintain this unqualified “compound” view of meaning we should need to be able to exclude from a logic of meaning those sentences (or statements or propositions)

---

<sup>183</sup> See also the discussion of this point in Goddard (1964, pp. 315ff).

that are meaningless, otherwise logical laws such as  $p \vdash p \vee q$  would have to be given up. —And in fact the metalogic of reference has precisely this purpose: to identify, eliminate, and thereby *exclude* projectively meaningless statements. We shall therefore accept the rule, when projectively meaningless statements are involved, that no compound statement in which they occur can possess meaning. In the context of our interests here, Moh's approach therefore is not of value.

Åqvist understood by a meaningless statement simply one that is neither true nor false. This, in itself, as I shall make clear, is an unwelcome restrictive constraint upon a logic of meaningfulness, for there are numerous varieties of expressions that we routinely accept as meaningful, and yet which are neither true nor false. To give examples acknowledged even in Åqvist's own work, he claims that meaningless sentences are to include sentences, e.g., that are imperatives, interjections, etc. (1962, p. 140n). —And yet, do we really wish to classify such sentences as "meaningless"? Surely they do possess meaning, though not of the truth-functional sort. Restricting meaning to truth-functional statements is undeniably limiting.

Åqvist proposed a minor variant of Kleene's three-valued logic, and used the third value to stand for "meaninglessness." The calculus  $\mathcal{A}$  that he formulated was intended to apply to philosophical problems of nonsense, but his system, like Moh's, suffers from a similar shortcoming: For Åqvist, the proposition "it is true that  $p$ " where  $p$  has the value "meaningless," is considered to be *false*. In the case of a projectively meaningless statement,  $\overline{p} = \mu$ , to claim that such a statement is true (or false) is itself without meaning.

*The Logic of Nonsense* (1949) by Sören Halldén (1923–2010) is a short 132-page monograph; it was published when he was 26, the year before he wrote a doctoral thesis about the concept of truth. His interest in the logic of meaningfulness (which he calls 'nonsense') was influenced by the work on criteria of meaning by Carnap, Schlick, Bridgman, etc. Parts of his book discuss verificationism, and touch only briefly on Bridgman's operationalism.

Halldén defines a meaningful proposition as one that is either true or false; propositions that are neither he regards as meaningless. But then he later adds the following curious comment:

[I]t may of course be the case that there exists some connection between the property of being true or false, and the concept of meaning. However, that is something to be proved. It must be stressed that there exists no directly discernible connection between the two concepts. (Halldén, 1949, p. 37)

He never went on to prove that there is indeed a connection between truth/falsity and meaning. His comment is made harder to understand when he remarked: “I want ... to distinguish the concept of intelligibility from the concept of meaningfulness. As far as I know, there may be intelligible propositions which are neither true nor false, and unintelligible ones which are true or false” (p. 38). In this way, Halldén implicitly suggests that the class of meaningful propositions may be larger than the class of propositions that are true or false (assuming that an “intelligible proposition” must be meaningful); he did not go on to provide an account of either “intelligibility” or “unintelligibility.”

He represents “truth” by means of the symbol ‘1’, “falsity” by ‘2’, and “meaningless” by ‘3’. His notation includes the following:

‘+’ the statement following this sign is meaningful, i.e., is true or false

‘-’ the statement following this sign is meaningless, i.e., is neither true nor false

‘~’ the statement following this sign is negated

He provides the following matrices for his basic connectives:

+	<i>p</i>
1	1
1	2
2	3

~	<i>p</i>
2	1
1	2
3	3

In the above table we see that the negation of a meaningless statement is meaningless.

<i>p . q</i>	1	2	3
1	1	2	3
2	2	2	3
3	3	3	3

If one of the conjuncts of a conjunction is meaningless, this makes the compound conjunction meaningless.<sup>184</sup>

–	$p$
2	1
2	2
1	3

We note that his ‘–’, “it is meaningless that,” is applied like an operator: When  $p$  is true,  $\neg p$  is false; when  $p$  is false,  $\neg p$  is also false; when  $p$  is meaningless,  $\neg p$  is true. And so he gives this definition:  $\neg p = \sim +p$ .

In a similar fashion, Halldén includes tables for  $\vee$ ,  $\supset$ , and  $\equiv$ .

Much of Halldén’s motivation in developing his logic of nonsense is summarized in his statement: “A theory of meaningfulness, a *logic of nonsense*, may make it possible for us to refrain from unprofitable attempts at the solution of problems having, because of their very nature, no scientific solution” (Halldén, 1949, p. 12).<sup>185</sup> He did not set himself the objective of constructing a rigorous deductive formal system, “believing such an attempt to be premature at the present stage. I do not believe that formalization is a necessary prerequisite for philosophical precision” (p. 22).

The paradoxical proposition, “This proposition is false,” is for Halldén meaningless because it is neither true nor false. And yet we should hesitate and realize that for us to know that this proposition is neither true nor false requires that it have “sufficient meaning” for the content which it expresses to be known to be neither true nor false. Halldén’s primary focus in his book relates to the logical paradoxes, and in his thinking about them he repeatedly applies his equation of meaningfulness with neither-true-nor-false, but he never recognizes that a certain level of putative meaning is required in each case in order for any of the paradoxes to be expressed in a manner with sufficient meaning that is recognizably paradox-generating.

Halldén’s approach leads to the following sample assertions:

<sup>184</sup> We note that he used the same matrices for negation and conjunction developed earlier by Bochvar, though Halldén was apparently unaware of Bochvar’s work at the time.

<sup>185</sup> In Halldén (1967/1961, p. 16), a book not translated into English, he discusses his related notion of “pseudo-problems”: “[T]he problems one poses may have no solution, may be questions without answers, pseudo-problems.... That a question is a pseudo-problem means that none of the sentences at issue is true. All the conceivable answers are incorrect.” (Translated into English by Furberg, 1973, p. 85).

$+p \supset +\sim p$       if  $p$  is meaningful (that is, true or false), then  
 $\sim p$  is meaningful (is true or false)

$+\sim p \supset +p$       if  $\sim p$  is meaningful, then  $p$  is also

From these it follows:

$+p \equiv +\sim p$        $p$  is meaningful (is true or false) if and only if  
 $\sim p$  is meaningful (is true or false)

$+p \equiv \sim -p$        $p$  is meaningful if and only if it is not  
 meaningless

He proves  $+\sim p$  (p. 47), i.e., it is meaningful to claim (it is true or false) that  $p$  is meaningless.<sup>186</sup>

- |     |                        |   |
|-----|------------------------|---|
| (1) | $++p$                  | basic thesis: it is meaningful (= T $\vee$ F)<br>to state that $p$ is meaningful (= T $\vee$ F) |
| (2) | $+p \supset +\sim p$   | from $+p \equiv +\sim p$  |
| (3) | $++p \supset +\sim +p$ | 2, subst. ‘ $+p$ ’ : ‘ $p$ ’  |
| (4) | $+\sim +p$             | 1, 3 MPP  |
| (5) | $+-p$                  | 4, def. ‘ $-$ ’   |

The expression ‘ $++p$ ’ deserves comment. It claims that it is meaningful to state that a proposition is meaningful. About this, Halldén remarks: “...it should be kept in mind that ... we allow ... for the possibility of ‘ $+$ ’ being systematically ambiguous. Then we may interpret the two occurrences of ‘ $+$ ’ ... as expressing two different concepts of meaningfulness” (p. 57). He doesn’t suggest what these two different concepts might be, but there is perhaps a parallel here to our earlier claim that iterated modal operators ( $\{7.6\}$ ) indicate distinguishable levels of meaning.

Writing as he did at a time when the verifiability criterion of meaning was already losing persuasive force among philosophers, Halldén expressed the wish for a more compelling criterion of meaning. Towards the end of his monograph he mentions in passing the concept of what he calls “designation,” which he does not define, giving only examples such as “ $a$  designates  $b$ ” and “ $a$  denotes  $b$ ” (p. 90). He saw in the concept of designation a possibly promising way to progress beyond the verifiability criterion:

---

<sup>186</sup> The proof given below is expanded somewhat to show omitted steps.

On the one hand the failure of the attempts at a satisfactory formulation of the principle of verifiability indicates that some new method of formulation must be found.... The theory concerning designation in particular seems to fulfill the intentions of the principle of verifiable [sic]. (p. 97).

This was perhaps an intuitive speculation on his part, one that seems to me (though I may be reading what I'd like to find here) to point to the potential value of a new and compelling principle of meaninglessness that could result from a well-formulated theory of designation/reference, but unfortunately Halldén did not develop this idea.

...

In the nearly three quarters of a century since Halldén published his all-too-brief book, the battle over competing criteria of meaning has, at least for now, subsided; arguments urging the acceptance and application of strict criteria of meaning now tend to be perceived as rare anachronisms; and, as a result, logicians today no longer take a lively interest in "the formal logic of meaninglessness." Patrick Suppes (1969/1950) was among the last philosophers with an interest in the formal logic of meaninglessness to publish about it. He made use of Bochvar's three-valued system (without, however, crediting Bochvar for his truth-matrices), to allow for a formal representation of operationally meaningless statements.

Some years later, logicians Leonard Goddard and Richard Routley (later Sylvan) published an important book relating to the logic of meaningfulness, *The Logic of Significance and Context* (Goddard & Routley, 1973). It may be fair to say that it remains the last substantial, book-length study of the formal logic of meaningfulness, one that is relevant to us here because their logic of significance attempted also to deal to some extent with meaninglessness.

Goddard and Routley set an ambitious goal for their work:

It is our purpose to develop a general formal theory of significance in terms of which significance claims, and arguments by means of which they are made, can be assessed. Thus we aim, eventually, to provide a logic, not previously developed, for much of modern philosophy" (Goddard & Routley, 1973, pp. 5-6).

A second volume was planned but unfortunately never published.

A number of observations and results by Goddard and Routley concerning the logic of meaninglessness that are relevant to the present study are worth mentioning:

They chose to make the concept of meaning (they call this the ‘concept of standard sense’) logically primitive (Goddard & Routley, 1973, p. 37). They did not give an explanation for this choice, but it is the same choice I have made and have explained earlier in light of the great range of varieties of meaning that we have recognized {11.1, Table B}. Goddard and Routley were familiar with Halldén’s book, which is mentioned fairly often, and they followed Halldén’s decision to regard sentences that are neither true nor false as meaningless (the term they use is ‘nonsignificant’). They claimed that if the fundamental principles of a significance logic can be stated meaningfully from a metatheoretical point of view, then they are formalizable, “and accordingly, from the point of view of the meta-metatheory, formalisable in a two-level formal system” (Goddard & Routley, 1973, p. 230). They aimed to develop their “significance logic” (perhaps extending to a group of such logics) in such a way as to *exclude* nonsignificant sentences, much as we have stated our intention to do; to accomplish this, they wished to formulate rules that would reject and bar meaningless sentences from the logic itself:

If a language is to be constructed in such a way that its rules exclude non-significant sentences, the actual construction must be carried out in a language in which such sentences appear. Rules for their exclusion cannot be formulated unless we can recognise them for what they are and make significant arguments about them. (Goddard & Routley, 1973, p. 233)

Goddard and Routley rejected the belief that to formalize a metalinguistic semantical system it would be necessary to adhere to a strict object language/metalinguage distinction (p. 236), a belief that was radical for its time.

They also adopted the principle, also followed in the present work, that the conventional negation of a meaningless sentence is meaningless:

The importance of this feature, namely that the classical negation of a non-significant sentence is nonsignificant whereas the negation of a false (logically false) sentence is true (necessary), is that we are thereby enabled to put aside from truth-valued and modal-valued consideration a whole class of pos-

sibly troublesome sentences, both the sentence and its negation and very many compounds of these. (p. 237)

Their approach led them to adopt the following matrices (p. 261), which by now should be familiar (they used ‘n’ for the third value to represent ‘nonsignificant’):

<b>&amp;</b>	t	f	n
t	t	f	n
f	f	f	n
n	n	n	n

v	t	f	n
t	t	t	n
f	t	f	n
n	n	n	n

~	
t	f
f	t
n	n

We note that these matrices are identical to those we have discussed by Bochvar, Halldén, and to those adopted in {11.4}, where the third value is represented by ‘ $\mu$ ’. Goddard and Routley adopted the principle, as we do in the present study, that if a portion of a compound statement is meaningless, this “infects” the entire statement, which then, considered as a whole, is meaningless.

Goddard and Routley followed Halldén in his definition of ‘+’ for meaningful statements, replacing this symbol with the connective ‘S’ which “will assign value false when what it applies to is nonsignificant and value true otherwise” (p. 246). They used ‘T’ to assign the value true, and ‘F’ to assign the value false.

How, then, did they define meaninglessness? This was their answer:

An assertion is nonsense, i.e. neither true nor false, when it exhibits significant incompatibilities, much as an assertion is logically false when it is false because of logical incompatibility. Nonsignificant sentences are not senseless, in the sense that they do not have senses or contents, and different non-



significant sentences may have different senses and different contents. (p. 249)

On the surface, the second statement in the above passage is a curious remark—that meaningless sentences are not without “sense” or “content.” What I understand Goddard and Routley to be suggesting, by implication and extension to the framework of the present study, is that even the special variety of meaningless statements that I call ‘projective’ expresses some “sense” or “content” which is such that it permits us to analyze such statements and show that they are indeed meaningless because self-undermining on the level of possible meaning. —Here, once again, the issue of putative meaning enters in, an issue that Goddard and Routley do not examine. A final comment about the quoted passage above: To be clear, the “significant incompatibilities” mentioned in the quotation are of course not associated by Goddard and Routley with the variety of metalogical self-referential inconsistency that comprises a central focus of our study.

Later in their book, the authors develop an axiomatic approach to nonsignificance logic. The theorems they derive include these examples:

- |                                      |  |
|--------------------------------------|--|
| $\vdash S \sim Sp$                   | it is significant (i.e., meaningful) to say that a proposition is not significant (theorem 17) |
| $\vdash \sim Sp \supset \sim Tp$ and | if $p$ is not significant, then $p$ is not true;   |
| $\vdash Tp \supset Sp$               | if $p$ is true, then it is significant (theorem 18)  |
| $\vdash \sim Sp \supset \sim Fp$ and | if $p$ is not significant, then $p$ is not false;  |
| $\vdash Fp \supset Sp$               | if $p$ is false, then $p$ is significant (theorem 19) (p. 393)                                 |

After the publication of Goddard’s and Routley’s book, there followed a long period without major contributions either to the formalized logic of meaningfulness, or to that of meaninglessness. Beginning in the 1990s, logicians such as Graham Priest, Richard Routley, J. Norman, JC Beall, B. Armour-Garb, Francesco Berto, J. Woods, and others,<sup>187</sup> who opened formal logic in ways that either tolerate or explicitly affirm contradictions, might be expected to deal with the logic of meaninglessness, but we instead find that so-called ‘paraconsistent’ and ‘dialetheic’ logics do not advance our under-

---

<sup>187</sup> Representative works include Priest (1987, 1995); Priest, Routley, & Norman (1989); Priest, Beall, & Armour-Garb (2004); Berto (2007); Woods (2003).

standing of the logic of meaninglessness beyond the truth-functional approach taken long ago by Bochvar, Halldén, Moh, and Åqvist. Priest (2006, p. 267), e.g., has continued to hold to the view that an understanding of meaning, and hence of meaninglessness, has to be made in terms of truth conditions: “I know of no other approach to meaning that I find satisfactory.” The present study develops such an approach, not in terms of truth conditions.

#### **12.4 Reflections on logics of meaning and meaninglessness**

The shared objective of the formalized logics of meaning and meaninglessness that I have discussed is to develop ways of dealing with meaningful statements and those that are devoid of meaning using the tools of formal logic. Beyond this broad goal, the logics most relevant to the present study are those that expressly wish systematically to exclude meaningless statements from a role in formalized logical reasoning. To this end, some criterion or set of criteria of meaning is necessary; through the application of such criteria, it is theoretically possible to restrict a given formal logic to only meaningful statements. However, none of the approaches I have described has proposed, let alone justified, such a criterion or set of criteria of meaning. As a result, the logics of meaning and of meaninglessness that have so far been developed have failed to supply us with a methodology to exclude meaningless statements. Goddard and Routley, for example, mention that the meaningless statements they would wish to exclude are those that “exhibit significant incompatibilities,” but we are left without the tools needed by means of which to identify, eliminate, and therefore exclude them.

Secondly, the logicians we have met in this section all share in the view that meaning is to be defined in terms of statements that are true or false. I have commented on the limiting nature of this restriction in light of the very broad range of varieties of meaning we have already taken note of. The logicians who have been interested in logics of meaningfulness and meaninglessness have all been philosophers, and as philosophers they have exhibited the characteristic twentieth-century logicians’ inclination to focus upon language, and, in particular, the language of declarative/indicative sentences. However, we do, of course, accept—even when we confine ourselves to meaningful uses of language—that meaningful sentences may be of other types, whether imperative, interrogative, exclamatory, etc. But my point here does not concern sentence types, but rather the nature of meaning.

### 12.5 The logical priority of the bonded pair reference-and-meaning over truth-functionality

In the case of indicative sentences that are true or false, their capacity to function as communicators of truth or falsity relies upon their logically prior capacity to identify, to varying degrees, that which they are about, in ways that are informative, which is to say meaningful. For this to happen, a statement must refer in some identifying manner, which informs us what the statement about. That a statement *must refer in some way that makes sense*, that conveys information, is logically prior to its being a candidate of truth or falsity.

I shall call identification/identifying reference and meaning a ‘bonded pair’: the relation between them is *functional* and *systemic* in the sense that each is a function of the other and neither is possible without the other; both presuppose relativity to some appropriate framework of reference; and both are rendered possible by the metalogical preconditions of reference of that framework’s constitutive structure.

It is evident that in order for a statement to be true or false, the statement must first, in the logically prior sense, have a meaning,<sup>188</sup> and for it to possess a meaning, the statement must be about something in the sense that it must communicate information. We commonly recognize this, for example, when we accept that to know whether an indicative statement is true or false, one must first know what it means—that is, what it is about, what the statement vaguely or specifically refers to. Meaning is built on and is ingredient in reference, but reference itself must make sense. It is a mistake (and indeed projective) to attempt to consider either without the other. They form, so to speak, a bonded pair.

As we noted earlier in {5}, pure ostensive reference—expressed, for example, by the use of a finger or arm to point—is probably one of the most ancient, anthropologically primitive forms of reference to which sounds and then words became linked. Yet, even in its basic form of simple pointing, meaning is unavoidably ingredient in and is built upon the very way in which pointing occurs—which may nonverbally express surprise (“Look at that!”), alarm (“Watch out!”), casual interest (“Just look over there.”), perhaps merely (“Direct your attention there.”), etc. Even simple pointing is a function of a context in terms of which it is able to identify and convey information. In considering such minimally simple ostensive reference, it might superficially

---

<sup>188</sup> Rescher (2014, p. 63) recognized this when he asserts in passing, but does not justify, the claim that “before there is truth there must be meaning,” but he makes no mention of the role of reference.

be said that reference/identification is, logically, first; that meaning is derivative; and that truth-and-falsity are later descendents. But, when carefully examined, reference and meaning form a bonded pair. Even in the simplest ostensive reference, meaning is unavoidable embedded.

Philosophers of language have studied ways in which agents use sentences to state things, with the result that an analysis of language use has become a central focus of philosophical interest. From the metalogical point of view of this study—that is, in which the dominant concern is with the preconditions of identifying reference—language use is logically posterior both to the satisfaction of such preconditions and to their associated framework relativity. The consequence we have been led to is that, in order to possess a meaning, any statement must be non-projective, as a necessary condition of its possible meaningfulness.

For each of the 18 varieties of meaning we have recognized (listed in Table B of the previous chapter), it would, at least in theory, be possible to formulate some set of criteria that must be satisfied in order for that form of meaning to be possible. Suppose that this has been done. Then failure to satisfy that set of criteria would render that particular variety of meaning without meaning: non-satisfaction of that set of criteria would result in a corresponding form of meaninglessness.

What is important to realize in this fictional project is that what we identify as “meaninglessness,” of whatever variety, derives *its* metatheoretic meaning from the non-satisfaction of a set of criteria. Meaninglessness presupposes that some conditions or criteria—usually unstated and implicit—have been applied.

The essential relationship to framework-relative metalogical preconditions of reference is presupposed in order for meaning of any kind to be possible, whatever the individual concrete properties of that relationship. Deny that relationship, reject it, undermine it, and you get, as a result, something that is without meaning.

The current and prevailing view of meaning continues, as in the last century, to be linguistic—that expressions and sentences made up of such expressions are either meaningful or meaningless, and that they are meaningful if they are true or false, and meaningless if neither. In the chapters that follow, the position we shall justify is that meaning, if it is to be possible at all in the conventional linguistic sense—or in any other sense—must first comply with preconditions of valid reference. Referential consistency, as defined in the previous chapter, will become a fundamentally necessary, but not sufficient, criterion of meaning. Referential consistency, in other words, will turn out to be a metalogical prerequisite of meaning, and hence a metalogical prerequisite

of the linguistically linked concepts of truth and falsity, which are logically posterior to it.

Due to the metalogical priority over truth-functionality of the bonded pair identification/identifying reference and meaning, we see that existing logics of meaningfulness have failed to focus on what is ultimately fundamental. Indispensably linked to the possibility of meaning is identification/identifying reference, not truth-functionality.

By employing referential consistency as criterion of meaning, we shall have a method to exclude from the resulting metalogic of reference forms of meaningfulness—of whatever variety, linguistic or otherwise—which we wish to identify and eliminate from rational consideration and discourse. We shall then possess a means to detect as well as eliminate a specific, widespread, and fundamentally important variety of what Goddard and Routley had called ‘significant incompatibility’, and a means that fulfills Halldén’s wish for a strongly compelling criterion of meaning, so that this pervasive form of incompatibility that results in meaningfulness can be excluded from the scope of rational thought and its expression. We will have available a generalized criterion of meaning, one which will enable us to identify, and then to eliminate, certain well-defined forms of self-referential inconsistency, and by doing this, will lead to a recognition, one which itself cannot *not* be accepted without referential incoherence, of the most general, invariant principles of reference that govern what can—in principle—be meaningful. In this way, we shall make an advance in developing Kant’s embryonic notion of a “negative science,” a *phaenomenologia generalis*, able to distinguish the meaningful from the meaningless in order to insure that only meaningful concepts, propositions, and statements remain for rational use.

# 13

## Projection

**I**t will be useful at this point to group together the following results reached in earlier chapters that we'll find relevant to the concept of projection:

- ◆ We recall that a primary interest of this study is to develop what I've called a 'philosophically neutral approach', that is, one capable of identifying universal, invariant principles that underlie and govern coherent, rational thought and discourse, analyzable from the standpoint of a level of maximum theoretical generality;
- ◆ A non-linguistic, non-relational conception of reference was introduced, one which—at least provisionally—avoids incorporating in the concept of reference such metalogically unanalyzed notions as linguistic agency, intention, volition, and causal action;
- ◆ From the proposed metatheoretical standpoint, a general theory of possibility was formulated in terms of which structural/systemic presuppositions and presuppositions of identification can be determined and employed in the context of a metalogical approach to transcendental self-referential argumentation;
- ◆ Framework relativity and associated framework-relative field theory were recognized as indispensable to the resulting metalogic, in terms of which the identification of particular objects reference is rendered possible;
- ◆ From this metalogical standpoint, an initial statement of the framework relativity of ontology was given whereby an object has ingredient in it, as an integral constituent of

its identity, the constitutive structure of the reference frame that forms the basis for its identifiability;

- ◆ Referential consistency was developed as a necessary, intrinsically determined general criterion of meaning, whether such meaning is linguistic or of other varieties;
- ◆ A solution to the problem of putative meaning was given and the related notion of delusions of meaningfulness was introduced;
- ◆ In examining formalized logics of meaning and of meaninglessness that have so far been formulated, we saw that none has proposed, and none has satisfactorily justified, a criterion or set of criteria that would provide a strongly compelling method to exclude meaningless statements; and
- ◆ We recognized the logical priority of the bonded pair, reference-and-meaning, over truth-functionality.

On this basis, we've come to see that any object of reference, that its very identity, is a function of the set of parametric constraints within which its identification can occur. Such constraints form, metaphorically, the interstices within which what is possible is determined: They form the network of positive and negative constraints that define possible objects of reference relative to a given reference frame. In this sense, such constraints function as structural/systemic presuppositions that must be satisfied in order for any variety of meaning to be possible.

The objective of the metalogic of reference, integrating the results summarized above, is to demonstrate that there exist inescapable "thresholds of meaningfulness," and to establish the boundaries of meaning that any conceptual/theoretical framework determines.

### 13.1 A therapy for concepts

Throughout philosophy's history, philosophers—and, in recent years, also some cognitive scientists—have sought to study certain dysfunctional ways in which people think and behave. Many philosophers have had a strong interest in fallacies, errors of inference and deduction, invalid arguments, and other

pitfalls that variously affect human thought, formal and informal logic, prudential, plausible, and probabilistic reasoning. Some of this work has been called, by the authors themselves, ‘therapeutic’.

Philosophical analysis considered as a kind of “therapy” has historically been approached in two significantly different ways. The first has sought to disentangle confusions that come about as a result of the *misuse of language*. This is the method recommended by Fritz Mauthner, Ludwig Wittgenstein, Brian Farrell, Gilbert Ryle, John Wisdom, and others. The second method has sought to undertake a rigorous analysis and treatment of confusions that are essentially conceptual, confusions that are due to a *misuse of ideas*—accompanied, of course, by flawed thinking on the part of those who implement such ideas. This second approach was first proposed and developed by the present author.

The two approaches are fundamentally distinct:

Wittgenstein sought to direct attention to ways in which, he believed, the misuse of ordinary language can bring about many of the fundamental confusions that occupy philosophers. His fragmentary, aphoristic writings pointed to a loosely defined set of analytical techniques that might help philosophers escape from the “fly-bottle” into which their inappropriate and misleading use of language trapped them. Brian Farrell (1946) variously called Wittgenstein’s approach a ‘semiotical psychiatry’ (pp. 135, 144), ‘semiotical therapy’ (p. 139), and ‘therapeutic positivism’ (pp. 143, 146-147). As expressed in Farrell’s commentary, Wittgenstein advanced “a psychiatric procedure for settling worries and conflicts of a certain type” (p. 35). In Farrell’s eyes, these philosophical worries and conflicts constitute a “disease” of a certain sort, one that comes about as a result of “linguistic confusion” (p. 141). In a later paper, Farrell (1990, p. 8) went so far as to recommend that such “semiotical psychiatry” should “be dovetailed into ordinary psychiatric treatment,” something which has never happened—very likely much to the relief of many non-philosophical mental health clinicians!

Suggestions that philosophy of ordinary language ought to be regarded as a genuine variety of “therapy” have tended to be impressionistic, metaphorical, and, in the present author’s view, overstated. Stephen Toulmin made use of the therapeutic analogy by coining a special work for the “diseases” of language misuse:

[P]hilosophical theories were [according to Wittgenstein] to be diagnosed as symptoms of misconceptions about our *everyday* language—“cerebroses” (so to say) comparable to the neuroses which spring from misconceptions about our



affective relationships. They were accordingly to be “treated” by a philosophical therapy adapted to the specific intellectual cramps of the individual patient. (Toulmin 1969, p. 60)

John Wisdom similarly used words “with a clinical flavor” in order to call attention to states of “philosophical stress” (Wisdom 1953; Dilman 1984, 1996). He, somewhat like Wittgenstein, sought to treat the peculiar linguistic confusions, the “mental cramps,” that beset philosophy.

The second approach, originated by the author, is fundamentally distinct. In the Introduction, I mentioned ‘*conceptual pathology*’ and ‘*conceptual therapy*’, terms which, beginning in the 1960s, I coined to refer, respectively, to certain self-referentially self-destructive *concepts* (and not simply misused or misleading *language*), and to a maximally general approach to conceptual analysis that can detect and eliminate such self-undermining concepts. In the latter sense, conceptual therapy specifically seeks to deal with self-undermining concepts and resulting claims, whether encountered in the belief systems of individuals, or in the conceptual structure of philosophical positions, scientific theories, or theories advanced by other disciplines. The focus of analytical attention is shifted away from the concrete uses to which individuals put language or express their thinking, focusing instead upon the referential presuppositions of whatever concepts and claims involving them that are subjected to analysis. To the extent that such an approach is regarded in any sense as comprising a “therapy,” it is a therapy for *concepts*, and not for *people*. —Some chapters back, we recognized a parallel distinction between pragmatic self-referential argumentation, directed at the ways in which individuals propound and defend arguments, and the metalogical variety, which is maximally general and can be used to study the preconditions of the possible meaningfulness of such arguments, formulated from the standpoint of metatheory, detached from particularizing and often psychologically rooted exchanges among individuals, who will often seek by whatever avenues may be expedient to defend their beliefs with the “almost inconceivable hardihood” remarked on by Passmore.

To make the above distinction clear, in {11.4}, for instance, we briefly considered Bridgman’s remarks about the contrived concept of “universal cosmic shrinkage.” There, it was the concept itself, the very idea, that the universe might be shrinking in a homogeneous and therefore undetectable manner which was the subject of Bridgman’s comments. Another quotation, in {12}, also from Bridgman, similarly directed attention to the concepts of absolute length and absolute simultaneity. The focus there, too, was upon concepts, not language use. In the same way, when in a later chapter we shall

consider the concepts of momentum and position in quantum theory, it will be an analysis of the concepts themselves, and not “ways in which we use language,” that will concern us.

Philosophical analysis with a direct and unapologetic focus upon *concepts*, rather than upon the ways in which human agents *use language*, has risen and fallen in popularity in rhythm with shifts in philosophical fashion, style, and favored paradigms of research. Some of these shifts allow for the translatability of one framework to another, but not all do.

Philosophers of language may be tempted to believe that any conceptually focused analysis can be translated into linguistic terms—as evidenced by the very fact that conceptually focused analyses find their expression in language, whether ordinary, technical, or mathematical. However, when a shift is sufficiently radical and revolutionary, changes in language use are not adequate either to recognize or to explain the transformation in the very thought processes of individuals who are closely affected by such a shift. This has been the case when we consider major conceptual revolutions that mark shifts, for example, from the Ptolemaic to the Copernican models of the universe, from the Copernican to the relativistic, or from pre-quantum to quantum physics. These conceptual shifts are not reducible simply to changes in the ways in which physicists choose to use language, but instead such shifts profoundly affect the ways in which physicists conceptualize and proceed with the evolution of their discipline. Such shifts affect not only human thought processes, but the design and development of instruments and experiments, and the hopes and expectations of researchers. These are not matters of language use.

My concern in developing conceptual therapy has been based on a recognition that the shift which comes about through the variety of metalogical analysis described in this book is sufficiently undermining of past ways of thinking about philosophical problems that this shift is not reducible to preferences in ways in which philosophers use language; the proposed shift has, I believe, the potential to change the way philosophers—those who are mentally prepared and willing—*conceive* of their work, and to change the *way* they undertake it. It will of course be up to readers to decide whether this is indeed the case for them.

### 13.2 The term ‘projection’

A brief discussion of the origins of the term ‘projection’ as we shall use it will cast light on the underlying human psychological predisposition that dovetails with and fosters the conceptual malfunction it refers to, while serving as an introductory explanation of that malfunction itself.

The term ‘projection’ has been used with a wide range of meanings in different disciplines and by different authors. Faced with a need for a specialized word for a new concept, an author is forced either to coin a word or to employ an existing word some of whose already established meanings bear resemblances to what is needed. I’ve chosen the second path for several reasons, two that are psychologically based, one that derives from mathematics, and one that is allegorical.

The word ‘projection’ commonly may be used to communicate the sense of actively *projecting* an object, of sending a *projectile* on its way, of *throwing* something. With an interest in both philosophy and psychology, I saw early in my life that many claims made by philosophers, as well as claims made by non-philosophers, appear to express a psychological propensity—often even a *willfulness*—to “throw” or to “over-extend” claims beyond the limits permitted by the very framework presupposed by such “throwing,” much as a ball player might wish to hit a ball out of the ballpark. This is of course colorful and impressionistic language, but it may help to capture one basic aspect of the meaning of ‘projection’ in this work.

Second, also originating from psychology, are several related meanings associated with the word ‘projection’ that come to us from Freud. He began to use the word relatively early, during the years 1894–1896, when he described projection as a psychological tendency to attribute to another person unpleasant or unacceptable emotions.<sup>189</sup> Some years later, he changed this characterization into the notion that projection is an explicitly defensive process in which a person attributes his or her own feelings to others in order to avoid the unpleasant self-awareness that they represent his or her own disliked and rejected propensities.<sup>190</sup> This is the most familiar sense of projection associated with Freud; we’ll call it ‘emotional projection’.

Later, Freud suggested, but did not develop, two other lesser-known meanings of the term ‘projection’ that are more relevant to the meaning I employ in this study: They are somewhat less affectively centered than emotional projection and come closer to having epistemological application. In Bartlett (2005, pp. 77ff), I differentiated these and named them ‘reifying projection’ and ‘rule-prescriptive projection’. The first of these Freud described as follows:

Under conditions whose nature has not yet been sufficiently established, internal perceptions of emotional and intellectual processes can be projected outwards in the same way as sense

---

<sup>189</sup> Freud (1953-74/1894, pp. 43-69) and Freud (1953-74/1896, pp. 159-188).

<sup>190</sup> Freud (1962/1912-13, p. 61).

perceptions; they are thus employed *for building up the external world*, though they should by rights remain part of the *internal world*. (Freud 1962/1912-13, p. 64, my emphasis)

I shall call a ‘reifying projection’ an inappropriate ascription of external existence to phenomena that are inherently internal.

In contrast, I shall call a ‘rule-prescriptive projection’ one that comes about in the following way:

[P]rimitive man transposed the structural conditions of his own mind into the external world.... The technique of animism, magic, reveals in the clearest and most unmistakable way an intention to impose *the laws governing mental life* upon real things....” (Freud 1962/1912-13, p. 91; my emphasis)

A rule-prescriptive projection expresses the belief that events in the real world conform to and obey the rules that structure inner experience. Like the two other varieties of projective belief, emotional and reifying, rule-prescriptive projection involves an unjustified extension of subjective interests.

The common psychological propensity to “throw” claims—to project or over-extend them—beyond the limits permitted by the reference frame(s) which such claims presuppose lies at the core of the concept of projection. Freud’s notions of reifying and rule-prescriptive projection especially exemplify this human propensity.

Beyond these origins of the term ‘projection’ as I use it in this study, for readers acquainted with projective geometry, there is the instructive mathematical meaning of the term ‘projection’. Stated very generally, projective geometry comprises a study of the properties of geometrical objects that remain invariant under projective transformations.<sup>191</sup> In projective geometry we encounter much the same emphasis upon a rigorously analytical self-consciousness of the role of coordinate frameworks and of coordinate transformations as is found in connection with the previous analysis of *perspectives* in {10.5}, where reference to projective geometry was also made.<sup>192</sup> There, we encountered the isomorphism and complementarity of

<sup>191</sup> Cf. Bartlett (1970, Vol. I, p. 187n; Vol. II, p. 124n).

<sup>192</sup> Work in projective geometry led, for example and coincidentally, to the first mathematically formalized expression of the *theory of perspective* in art, demonstrating that lines that are parallel meet at a vanishing point at infinity, and therefore should be drawn accordingly, so that if extended they meet at a point.

frameworks, concepts associated with the invariance of properties of geometrical objects under projective transformations.

Finally, as to the allegorical origin of the term ‘projection’ as used in the present study: Many readers are familiar with Plato’s allegory of the cave in which he portrayed a cavern in which normal, philosophically unenlightened humanity sits facing a wall upon which are “projected” the shadows of objects belonging to the world of genuine reality. The projected shadows that populate mankind’s wall-directed perception are mistakenly equated with what is ultimately real. We shall find that the human propensity to fall victim to the remarkable conceptual malfunction that mistakes projections for genuine, meaningful reference can be understood in much the same allegorical-metaphorical sense. As we shall see in coming chapters, metalogical projections—in the over-extended sense that surpasses the possibility of meaning—comprise Reality for much of humanity.

### 13.3 Projection in relation to other forms of meaninglessness

Expressed in an abbreviated and basic form, a projection occurs when we believe that that to which there is reference is autonomous of the very conditions that render such reference possible.<sup>193</sup> This clearly is not Freud’s emotional, reifying, or rule-governed projection, nor is it projection in the geometric sense. It does involve, however, the projective sense in which one may attempt to “throw” or “over-extend” the legitimacy and meaning of a claim beyond the limits prescribed by the preconditions of the “throwing” itself. Let us consider this idea with a little more care.

We identified some 18 varieties of meaning in {11.1}, along with eight of the possible forms that theories of meaning may take. There may well be more than these, in both instances, but the evident plurality and diversity of kinds of meaning should make it clear that there are likely also to be many ways in which meaninglessness comes about and may be investigated—not only projections as studied by the metalogic of reference.

Philosophical claims, positions, and theories (as well as their frequent counterparts in other disciplines), however, are usually expressed or can be reduced to indicative, declarative statements—i.e., assertions—that claim to be true, or that claim that other statements—often of course those proposed by other philosophers with whom one disagrees—are false, inconsistent, incomplete, misleading, etc. Or sometimes, it is argued, if these other statements

---

<sup>193</sup> Earlier works by the author that develop and apply this concept include Bartlett (1971; 1975; 1976; 1982; 1983; 2005, Part II; 2011, Chapters 2, 8).

that one rejects are not false, etc., then they are labeled indeterminate in meaning, or simply devoid of sense, as is sometimes said in the case of paradox-generating statements. And then, in addition to these various values which statements may have, we now also include projective statements, whose value we've expressed by ' $\mu$ '.

If we look at the bigger picture as it concerns varieties of meaning and meaninglessness, we are forced to admit that the class of possible statements includes a large variety of different sorts of statements beyond those that can be represented in only a three-valued logic, as was employed in {11.4}. The significant range of statement values must, if we are to broaden our scope of recognition, allow for *multiple values* and clearly not be limited to statements whose values are either T or F. The same is also the case in connection with the range of meaningless statements, if we are to recognize that their diversity requires more than a single value to represent meaninglessness.

For this reason, the significant range was left open in {11.4} to allow for statements with such values as indeterminacy, probability, etc. We refrained from limiting meaningful statements to those that are true or false only. But we nonetheless made the assumption for the sake of simplicity that the range of statement values that we wish to deal with requires only three values: T, F, and  $\mu$ . This was a convenient assumption to make given that the main formalized approaches to the logic of meaning which have been proposed are three-valued; as we saw in the last chapter, they endorse the view that meaningful statements are true or false, and that those which are neither, are meaningless (or undefined, or indeterminate).

For the specific objectives of this study, we may continue to make this assumption, as long as we remind ourselves that it is an assumption made purely to simplify matters, but not oversimplify them. Since our primary interest in the metalogic of reference is to be able to identify and eliminate projective concepts and statements using them, a three-valued logic with values T, F, and  $\mu$  suffices; but we should keep in mind that there is much we therefore omit, both in terms of varieties of meaningful statements, as well as varieties of those devoid of meaning.

### 13.4 Historical intimations of the concept of projection

PHILONOUS. [C]an any more be required to prove the absolute impossibility of a thing, than the proving it impossible in every particular sense that either you or any one else understands it in?

...

[T]o assert that which is inconceivable is to talk nonsense: is it not?

...

My business was only to shew, you meant *nothing*; and this you were brought to own. So that in all your various senses, you have been shewed either to mean nothing at all, or if anything, an absurdity. And if this be not sufficient to prove the impossibility of a thing, I desire you will let me know what is.

– George Berkeley (1999/1710-1713, pp. 167, 155, 168)

When we look back at the history of philosophy, we find only a few philosophers who seem to have touched on the margins of what I call the ‘concept of projection’. But if at all, their “touch” has been ambiguous, indirect, or tentative and fleeting. In this section, I’ve selectively chosen from philosophy’s history, for brief comment, several philosophers whose works, at least as I read them, bear certain differing but relevant relations to the concept of projection. There are other philosophers who could be added to the small group I’ve chosen, but the following sampling is intended only to illustrate the general philosophical motivation behind my development of the concept of projection. In this, I’m not concerned with the textual accuracy of my interpretations, which are intended here not as scholarly commentary but rather to explain the rationale that led to the concept of projection.

Irish philosopher Bishop George Berkeley (1685-1753) can serve as a first example: He wished to prove—as translated and re-expressed sympathetically (perhaps over-sympathetically) in the framework of this study—that what we are capable of *meaning* by “external reality” is necessarily mind-relative,<sup>194</sup> since, in his view, it isn’t possible to refer to external reality without presupposing a mind that provides the reference frame necessary for this to be possible. To assert otherwise is, for Berkeley, “to assert the inconceivable” or simply meaningless; it “is to talk nonsense.” In suggesting this translation, I take it that the concept of what is “possible” is functioning in a central way in Berkeley’s position, that he, when he spoke of what is “conceivable,” really had in view, or intended to have in view, what we now call ‘preconditions of possibility’. This admittedly may be stretching a good deal to reach for a point of contact between Berkeley’s thought and my own. There

---

<sup>194</sup> As Berkeley (1999/1710-1713, p. 196) put it, “a relative existence, with respect to created minds.”

are, of course, those who interpret Berkeley in a contrary way.<sup>195</sup> We cannot settle this matter here, and can only say that if Berkeley “touched” the margins of the concept of projection, this touch was ambiguous.

More than a century and a half later, mathematician and philosopher William Kingdon Clifford (1845-1879)<sup>196</sup> made another but still elusive contact with the concept. It is not the concept as defined and employed here, but Clifford’s concept, for epistemological reasons perhaps related to ours here, led him to use colloquial, figurative, evocative language much as I have earlier in caricaturing the “throwing” of a claim beyond its possible meaningfulness.

In analyzing the basis for the human belief in other minds, he coined the noun ‘*eject*’ (from the Latin *eiectus*, or “thrown out”):

[T]he inferred existence of your feelings, of objective groupings among them similar to those among my feelings, and of a subjective order in many respects analogous to my own, — these inferred existences are in the very act of inference *thrown out* of my consciousness, recognised as outside of it, as *not* being a part of me. I propose, accordingly, to call these inferred existences *ejects*, things thrown out of *my* consciousness, to distinguish them from *objects*, things presented in my consciousness, phenomena.... How this inference is justified, how consciousness can testify to the existence of anything outside of itself, I do not pretend to say.... (Clifford, 1878, p. 58 )

Clifford proceeded to use the term ‘*ejects*’ to mean the contents of other minds, and developed the notion of “*ejective facts*,” which also have a wholly inferential status, to construct the idea of an objective, shared physical world. It is important to note that Clifford offered no justification for the belief in the inferred existence of *ejects*. Were *ejects* to be considered comparable to projections, in the specialized meaning used here, then no such justification—

---

<sup>195</sup> Priest (1996, pp. 481, 483) is a good example: “I take it that the ‘possible’ is doing no real work here.... Berkeley, like many people, thinks of ‘conceive to be possible’ as a simple equivalent of ‘conceive’.... Berkeley identifies conceiving with conceiving to be *possible*.”

Maybe, but I don’t think so. I think that here Priest has missed what is central to Berkeley’s position, which may stand or be undermined by which interpretation is correct. This is not a book about Berkeley, so I must leave this unresolved.

<sup>196</sup> Clifford’s work in mathematics is famous (e.g., Clifford algebra is named after him): To mention one example, he anticipated Einstein’s general theory of relativity by some 40 years by suggesting that gravity is a property of curved space. He was a brilliant man; sadly, he died very young at the age of 33.



in principle—would be possible, as the present study makes clear.

Decades later, William James (1842-1910) alluded to “ejects” in passing, without referencing Clifford:<sup>197</sup> “I speak not merely of our ideas of imperceptibles like ether-waves or dissociated ‘ions,’ or of ‘ejects’ like the contents of our neighbors’ minds” (James, 1904, p. 68). Concerning these “ejects,” he says “we and the object ... can *never* get face to face, *as in the case of ejects...*” (p. 73, italics added). James failed to explain just what this “can never get face-to-face” really means—in terms of the possible meaning of ejects, and therefore their possible justification. If indeed “ejects” exceed the capacity of any reference frame’s capacity to identify them with the extreme degree of framework autonomy that they purport, by definition, to have, then “ejects” would come closer to the concept of projection.

Like James, British ethologist and psychologist Conwy Lloyd Morgan (1852-1936) was influenced by Clifford’s notion of ejects.<sup>198</sup> Morgan used the notion as a critical standard to dismiss and ridicule the tendency of his zoologist contemporaries—for example, Romanes (1883, 1895/1883)—to attribute conscious states to animals, *or* to other human beings.<sup>199</sup> Morgan claimed of the attribution of consciousness to another that “it is an eject, an image of my own consciousness which I *throw out* from my self.”<sup>200</sup>

At the very end of the 1800s, American philosopher Josiah Royce (1855-1916), edged somewhat closer to the concept that I call ‘projection’. Here are several relevant passages from his attack on the doctrine of realism; I have italicized some of his phrases that could be interpreted as compatible with the concerns of the metalogic of reference:

---

<sup>197</sup> James seems also to have taken from Clifford the basic idea of stream of consciousness as described in James (1904), again without citing Clifford (for comparison, see esp. Clifford (1878, p. 63)).

<sup>198</sup> American sociologist Lester Frank Ward (1841-1913) was similarly influenced by Clifford, commending the value of the idea of ejects “as an expression of a truth that is somewhat difficult to grasp, and one that is broader even than Clifford himself supposed” (Ward, 1907, p. 423).

<sup>199</sup> For example, Romanes (1895/1883, p. 22) wrote: “The evidence derived from ejects is practically regarded as good in the case of mental organizations inferred to be closely analogous to our own.” In Romanes (1884, p. 380), he also used Clifford’s term: “[M]y knowledge of another human mind is no less ejective than is my knowledge of a dog’s mind....”

<sup>200</sup> From an unpublished manuscript by Morgan, quoted in Richards (1989, p. 378, italics added).

Elsewhere he wrote: “My neighbour’s mind is to me neither subject nor object; it is an eject thrown out from myself. Into every man that I meet I breathe an image of my own mind, and thence forth he becomes for me a living soul” (Morgan, 1892/1885, p. 22). “My neighbour’s mind, feelings, motions are ejects to me; they can never be objects” (p. 267).

The realistic theory ... by its own consequences, and just because its real objects are totally independent of its ideas, has nothing to do with any independently real object, and has no relation to the independent external world that its own account defines. Nor *can it ever come to get such a relation*. No realist, as he himself now must consistently maintain, either knows any independent being, or has ever, in idea, found himself related to one, or has ever made any reference to such a being, or has ever formed or expressed an opinion regarding one, or, in his own sense of the word "real," really believes that there is one....

[W]hat then is left us, if the realistic definition of Being, simply and rigidly applied, destroys its own entire realm, *denies its own presuppositions*, and shows us as its one unquestionable domain the *meaningless* wilderness of absolute Nothingness. Where, then, is *our* real world?... [W]hat we now learn is that any definition of absolutely independent being, being that could change or vanish without any result whatever for their fellows is, in all regions of the universe, natural or spiritual, a hopeless contradiction....

You ask him [the realist] to show you an Independent Being. He points at the table or at the stars. But those, for you, and for him alike, are empirical objects, bound up in the context of experience. Nor could *any possible* enlargement of experience *ever* show anybody a Being wholly independent. The only way to judge Realism, since experience is thus abandoned by the realist, is to examine the inner consistency or inconsistency of realistic doctrine. And we have seen that Realism is wholly inconsistent....

[Y]our assertion that the world is, involves a judgment that your present experience is interwoven in the whole context of the realm of valid or of possible experience. *This context*, however, *is not independent of its own fragments*.... And if you attempt to assert the Being of things in any more independent sense than this, *you struggle in vain to articulate your meaning*. (Royce, 1959/1899, pp. 136, 137-138, 245, 248-249).

Passages like those above touch on the margins of the concept of projection less hesitatingly. In these passages, Royce appears to express an implicit

recognition of framework relativity, and perhaps, by my sympathetic inference, also a recognition of the need to study philosophical claims in terms both of the referential presuppositions on which they rest, and of the relation such presuppositions bear to the possibility of meaning. At the same time, there are also traces of transcendental argumentation in his claims, as when he seems to allude to the constraints of possibility. But, like so many philosophical positions and claims, his assertions appear to stand fairly firm until one wishes for a full explanation of their justification, which he did not supply.

Rudolf Carnap (1891-1970) has been mentioned in earlier chapters in connection with criteria of meaning which he urged. Carnap's relevance to the present study relates largely to his methodologically self-conscious emphasis upon the role of linguistic frameworks in providing the contexts only in terms of which questions can, in principle, be answered meaningfully. Questions regarding the existence of objects can, in his view, then be answered meaningfully only in the context of an appropriate, accepted linguistic framework. Carnap called these meaningful questions '*internal questions*'. Questions about "the existence or reality of the system of entities as a whole," called '*external questions*' are "devoid of cognitive content"—that is, such questions pose "pseudoproblems" (*Scheinprobleme*)—and are meaningless (Carnap, 1950, pp. 206, 212; 1967/1928, pp. 301-343). For Carnap, "if an (ostensible) statement does not express a (conceivable) state of affairs, then it has no meaning; it is only apparently a statement" (Carnap 1950, p. 325). "A (pseudo) statement which cannot in principle be supported by an experience, and which therefore does not have any factual content would not be a statement, but only a conglomeration of meaningless marks or noises" (p. 328). This is what 'meaning' for Carnap means, and, as a result, the following is the meaning he associates with 'reality':

To be real in the scientific sense means to be an element of [a linguistic] system; hence this concept cannot be meaningfully applied to the system itself.... An alleged statement of the reality of the system of entities is a pseudo-statement without cognitive content. (p. 207)

Like Royce, Carnap claimed that realism is meaningless, but he extended the same criticism to idealism as well, which he argued is equally meaningless, and for the same reason:

*[N]either the thesis of realism that the external world is real, nor that of idealism that the external world is not real can be*

*considered scientifically meaningful.* This does not mean that the two theses are false; rather, they have no meaning at all so that the question of their truth and falsity cannot even be posed. (Carnap, 1967/1928, p. 334)

As he expressed this in passing, if realism and idealism are considered to be epistemologically divergent theses, it is “*only because of a transgression of their proper boundaries*” (p. 286, his italics)—a claim that would seem to make close contact with our present study, but it is a claim for which Carnap provided no further explanation.

Carnap was strongly committed to the above assertions, but when we search in his writing for their justification what we find are *re-assertions* of his endorsements of linguistic framework relativity and of his proposed criteria of meaning, yet we do not find a logically compelling justification for either claim. He clearly felt that both assertions should be convincing to anyone with commitments to scientific rigor and to the requirement that assertions be based on evidence, but the strict justification we would wish for is absent.

In previous pages, I have quoted a few passages from works by physicist and philosopher P. W. Bridgman (1882-1961). As we have seen, he urged, like Carnap, that we accept limitations upon that which we accept as meaningful—in Bridgman’s case, in compliance with his criterion of operationalism. Like Berkeley, Royce, and Carnap, Bridgman prescribed a strong dose of awareness of what I’ve called ‘framework relativity’ (he didn’t use this phrase), and he judged it meaningless when we seek to over-extend our claims through the belief that they are valid beyond the frameworks in terms of which they are formulated and justified. We may recall a previous quotation from Bridgman (1961/1927, pp. 28-29): “[O]ne is making a significant statement about his subject in stating that a certain question is meaningless.” He posed the question: “To what extent is what the instrument gives us colored by the instrument itself, or is the instrument capable of revealing to us something ‘independent of the instrument’?... [T]he concept of object, in and for itself, becomes meaningless” (Bridgman, 1959, p. 150). In discussing electrostatic fields, for example, he observed: “the fact remains that there is no instrumental method of giving meaning to ‘field-in-the-absence-of-the-instrument-of-measurement’” (p. 152). (He of course should have said that, *in principle* and in connection with such fields, there *logically cannot be* any such instrumental method.)<sup>201</sup>

---

<sup>201</sup> Coincidentally, Bridgman used the term ‘projection’, but in a sense radically different from its meaning in this study, and more akin to Clifford’s notion of ejects: By ‘projection’ Bridg-

Bridgman's conception of meaning is ultimately a recognition that all that we are capable of experiencing, of observing, of making meaningful claims about, is inextricably tied to ourselves and whatever observing apparatuses we employ. As quoted in part earlier:

The ultimate instrument is ourselves.... In general, we should never think of the world around us without also thinking of the nervous machinery in our heads by which we acquire knowledge of the world. To discover the best way of holding ourselves to this awareness constitutes what seems to me to be perhaps our most pressing intellectual problem.... [T]he world is not to be dissociated from the knower of the world.... [T]he object of knowledge is not to be separated from the instrument of knowledge. (Bridgman, 1959, pp. 153-154, 123, 169)

These claims that involve phrases like 'is not to be dissociated from' and 'is not to be separated from' are *injunctions*—assertions of rules of meaningfulness—which, according to Bridgman, we *ought* to follow if our thought, use of language, and behavior (at least as scientists and as scientific philosophers) are to be meaningful. If his operational criterion of meaning is accepted, for Bridgman it follows that when attempts are made to “dissociate” or “separate” the object of knowledge from the knower, meaningless statements result. But, as has been the fate of other traditionally proposed criteria of meaning, Bridgman did not prove that his criterion *must* be accepted on pain of incoherence.

...

This section has given a brief overview to illustrate points of possible contact between the thought of a small group of researchers and the objectives served by the concept of projection as developed in this work. We've seen that the psychological impetus that we find in projections to over-extend claims beyond the frameworks of their possible meaningfulness was to some extent hinted at by Clifford in his notion of ejects, whose inferred existence comes

---

man meant “the operation by which I give meaning to your use of ‘conscious’.... I shall call it the operation of ‘projection’; I ‘project’ myself into your position, that is, I imagine myself in your position, and I ask myself what I would be saying or doing in such a position” (Bridgman, 1959, p. 220).

Yet, at the same time, he accepted that there is no operational meaning that can be given to the question whether another person is conscious in the same way as I am (p. 223).

about in a thought process that he likened to “throwing something out and beyond one’s consciousness.” He did not seek to evaluate and criticize, much less condemn, that psychologically based process, nor did James later. Morgan, however, judged the human propensity to eject beyond one’s own consciousness in a critical light, much like later philosophers would employ criteria of meaning; ejects became for him an illegitimate and unjustified practice.

Other philosophers mentioned in this section were concerned to lay down guidelines and limits beyond which assertions become meaningless. With Berkeley, conceivability served as such a limit; with Royce, we seem to find an implicit affirmation of framework relativity, and also—with some stretching—a tacit recognition that the validity of philosophical claims requires an analysis of referential presuppositions and their transcendental relationship to the possibility of meaning. With Carnap, we find a more explicit endorsement of framework relativity in connection with the framework of any given language system, and the application of a standard of meaning for which he urged acceptance. Bridgman affirmed framework relativity with respect to that which we are capable of experiencing, so that for him the class of meaningful statements is limited by our capacities to experience, whether directly or by instrumental means. He urged acceptance of an operational criterion of meaning, from which it follows that attempts to “dissociate” or “separate” an object of knowledge from the framework of the knower result in meaninglessness.

In connection with all of these contentions that seek to establish limits of meaningfulness, endorsed and advocated by the sample thinkers considered in this section, no strongly compelling justification to back up these claims was given by any of them. Dispassionately considered, their injunctions appear to express restrictive philosophical-epistemological *preferences*, but such injunctions suffer from the weakness of unjustified maxims to which we are *asked* to consent and comply in logical, rational, philosophically and scientifically acceptable discourse.

As the reader is by now aware, the metalogic of reference seeks to remedy these shortcomings by demonstrating, first, that framework relativity is an indissociable property of any object in relation to appropriate reference frames in terms of which its identification/identifying reference is possible (and therefore not only linguistic frameworks), and, second, that referential consistency is a criterion of meaning that cannot, without metalogical self-referential inconsistency, be denied.

### 13.5 The concept of projection

Towards the beginning of this chapter, the choice in this study to focus analysis on *concepts* rather than *language use* was explained. It is in the context of that choice that I now want to consider the concept of projection in greater detail. What I shall have specifically in view is a manner of thinking or a way of conceptualizing, a pattern that describes how certain elements of thought are assembled and then mistakenly taken to be meaningful. There are two ways in which this could be done: by studying how people think in the specified erroneous way, or by examining the pattern itself, considered as an object of study, separated from the concrete, particularizing psychology of individual minds. To do the first is to engage in psychology; to do the second, as we shall see in this and later chapters, is to examine a central issue belonging to epistemology. It is in the second sense that in this section we shall study the concept of projection.

In {11.4}, a brief formalized description of projection was given. We recall that projections were symbolized there using vector notation, consisting of an arrow over a symbolized projective sentence. As understood in its general mathematical/physical sense, a vector (often termed a Euclidean or spatial vector) is an abstract representation of a relation having both magnitude and direction. A vector is frequently depicted graphically as an arrow (or directed line segment) drawn so as to connect an initial point and a terminal point; the vector is conceived of as the “carrier” from the initial to the end point (from the Latin ‘*vector*’, meaning ‘carrier’).

Projections have, so to speak, two distinguishable aspects, which in the first paragraph of this section was already mentioned: On the one hand, there is the psychological component of a projection when asserted by a particular mind, and, on the other hand, there is its epistemological component, when considered reflectively as a pure object of study.

When projections are asserted by individuals, the belief they invest in their delusions of meaningfulness can, in a rough and approximate sense, be quantified: Some projective assertions are adhered to very strongly, others much less so; they can often be ranked in terms of the comparative degrees of belief invested in them. The degree of belief invested in a projection may be regarded as the *magnitude* of its vector, which “carries” the putative reference in a certain direction.

The *direction* of a projective vector—and here I am forced to speak metaphorically and, as we shall see later, inaccurately—is always “outward,” that is, “out of bounds,” “beyond the limits of possible meaningfulness.” Projections purport to refer “outwardly,” they have this intended “direction.”

These two conceptualizable aspects of a projection express both direction

and magnitude, for which, when formalized, vector notation seems most suitable.<sup>202</sup>

It will be useful here if we re-state informally the formalized description of projections given earlier:

Let the term ‘object’ stand for any uniquely or vaguely identifiable object of reference, whether a real, fictitious, or abstract object, a state of affairs, an event, etc. We recall that a statement (sentence or proposition) is considered to be metalogically self-referential when the statement putatively implies that reference obtains to some object at a certain time and place, and that such reference in turn entails one or more preconditions of reference that must be satisfied in order for the statement, in principle, to have an acceptable significant value (true or false, but we, as noted, may also allow such values as indeterminate, probable to a certain degree, uncertain, etc.). For such preconditions to be analyzed, the statement evidently must be considered from a metatheoretical point of view, since not only what the statement purportedly asserts but what it referentially presupposes need both to be considered; the phrase ‘metalogically self-referential’ serves to inform us of the fact that we are situating analysis on a metatheoretical level.

When a statement is such that it denies one or more preconditions that must be met in order for it possibly to possess a significant value, then that statement is *projective*, or equivalently, it comprises a *projection*.

We have recognized that by no means all meaningless statements are projective. A projection involves a unique variety of invalid reference that must satisfy the following pair of conditions:

(1) *A projection requires as a condition of its possibility that a particular object of reference (hereafter, simply called a ‘particular’) be putatively disconnected from certain of its essential relations to the identification framework entailed by its possibility. In other words, there must be a purported severing of the essential relativity of the object of reference to its presupposed framework of reference.*

(2) *The particular must implicitly or explicitly be alleged to be in certain respects autonomous of its context of reference. Reference to the particular must be such that it at least implicitly denies or flouts one or more of the metalogical pre-*

---

<sup>202</sup> In Bartlett (1970, esp. Section 1.2, “The Concept of Vector”), I developed the concept of vector in connection with a transcendental study of phenomenological constitution. The concept of vector as studied there is fundamentally different from its meaning in the present book.



*conditions of the particular's identifiability. The reference is projective in these respects.*

These two conditions must be satisfied by any projection from the standpoint of a metatheoretical analysis of its presuppositional structure. We shall find that it is not usually the case for a projection to take the form of an explicit repudiation of the metalogical constraints of its own putative reference, although this may be so in deliberately constructed examples. However, condition (1) must explicitly be satisfied when the projection is expressed, according to condition (2), in the form of a claim of autonomy with respect to the presupposed framework of reference.

Condition (1) stipulates that a particular is to be considered as though it were dislocated from the reference frame necessary for its identifiability. Condition (2) stipulates that this putative dislocation is to be formulated in the form of an at least implicit claim that the particular is autonomous of its presupposed framework of reference.

Satisfaction of the first condition results in a denial of the possibility of the particular, a denial that comes about when the particular putatively is separated from the reference frame conditioning its possibility as an object of reference. Such a "separation" is, strictly speaking, impossible—as long as reference is actually understood as purportedly referring to a particular object of reference. However, what is claimed of that object of reference—that it is autonomous of the identification system to which its possibility is relative—constitutes a metalogically self-referential inconsistency, which we've called 'projection'. From the standpoint of the metalogical analysis of projections, the alleged autonomy of an object of reference relative to its presupposed framework of reference comprises a special variety of fallacious, self-undermining reference.

Condition (2) involves reference to the particular while the grounds for its possible identity and identifiability are ruled out. In this way, a projection attempts to sever a relation of relativity—the relativity of a particular to appropriate identification frameworks in terms of which it can be identified.

The conditions I've just described which, when met, determine a projection, are not complex, but they can be misleading. A given projective statement initially purports to refer to some object; that it refers to that object is implied by the statement; and yet we find, after reflective analysis, that the statement undermines the possibility that it could refer as it purports to do. Here, the problem of putative meaning, which we studied in the last chapter, enters in. We need to differentiate very clearly between *the pre-analytical, purported meaning* of the statement, and *its self-undermining nature* disclosed

*subsequent to reflective metatheoretical analysis.* The statement is in the following sense deceptive or illusory: Considered at face value, it claims to refer, but when looked at more closely, we find that it undermines its own capacity to do this.

Stated more precisely, the statement, if it is to be at all significant, if it is to have a value in the range of accepted significant values, must be *capable* of referring in the sense which the statement appears to communicate. To be capable of referring, we see, upon analysis, that certain preconditions must be met, and we see that the statement implicitly or explicitly denies one or more of these preconditions. We come to see, in other words, that the statement undermines its own *possibility* of referring, and not merely that it denies that it refers. This is important, for *it situates the projective self-referential inconsistency on a level on which the concept of possibility plays a central role; it is the modal level required by transcendental argumentation.*

When these conditions are met by a projective sentence, the natural tendency is often to believe that, since the statement undermines its own capacity to refer, it would be valid to infer the negation of the statement. But we recall that the negation of a meaningless statement is no less meaningless. Instead, as long as we maintain a central interest in detecting and eliminating meaningless statements, sentences, or propositions, detection of a projection calls for its *rejection, not the affirmation of its negation.* By rejecting a projection, we obtain a statement that is self-validating—a statement that cannot itself be denied without metalogical self-referential inconsistency.

As we follow the above heuristic guidelines and establish a corresponding way of thinking we shall find that we gradually reach a delineation of the “thresholds of meaningfulness” associated with any particular reference frame. Mapping these thresholds of meaning is one of the main tasks of the metalogic of reference.

We have seen how, for any given frame of reference, one or more fields of possible objects are organized as a function of prerequisites that such objects must satisfy as possible objects of reference relative to that framework. These metalogical prerequisites or preconditions serve as a network that delimits *possible* candidates for reference, while, at the same time, ruling out or excluding reference to what does not comply with those preconditions. It is in this sense that such metalogical preconditions of reference determine what is possible relative to a given reference frame, and what is impossible.

The implementation of the heuristic guidelines that define the metalogic of reference as a form of conceptual therapy ideally leads to what in this work may be called a ‘*critique of impure reason*’: In enabling us to recognize the preconditions of possible reference, its task is negative in the sense of Kant’s

“*negative science*” as discussed in the preceding chapter. If we are able to determine the most fundamental restrictions on valid reference, we will have negatively defined the domain of possible sense. From this metatheoretical standpoint, the metalogic of reference makes it possible for us to recognize the fundamental constraints upon possible reference, the boundaries that define what is referentially forbidden—that is, the limits beyond which reference becomes meaningless.

The concept of projection may most clearly be grasped in the sense in which reference to objects, as described in previous chapters, is understood in terms of identification from the standpoint of a coordinate frame. Understood in this way, projective reference is reference which does not conform to, and which implicitly or explicitly conflicts with, the system of constraints that provides the basis for such putative reference from the standpoint of that coordinate framework. Projective reference is a putative coordination that is invalid because the purported reference conflicts with what must be granted for it to be possible. When reflective analysis reveals this, the meaninglessness of the projection is made evident.

In short, when reflectively analyzed, a projection is seen metalogically to entail an unsound and self-undermining form of reference. A projection comprises a form of invalid reference which, were we to use Kantian language, conflicts with its own “transcendental, presuppositional structure”: When one considers preconditions of reference in the sense in which they functionally express the concept of possibility, then a projection may be thought of as a transcendently self-abnegating reference. A projection, in other words, constitutes *invalid reference on the level of possibility*.

For simplicity and expedience in communication, we have defined ‘projection’ in terms of statements (sentences or propositions) that meet certain conditions. Yet, our principal interest remains an analysis of *concepts* in terms of their metalogical referential—i.e., structural/ systemic—presuppositions. In Part III of this study, we shall put the concept of projection to use in this way. Our analytical focus will be upon certain concepts, described of course in language but not reducible to language use, concepts that can be associated with characterizable patterns of thought which are at once self-undermining on the level of possible meaning and which form the basis for delusions of meaningfulness.

In an analysis of projective concepts undertaken by the metalogic of reference, we shall find that we’re confronted not merely with actual misguided, delusional uses of particular concepts, but with concepts which preclude their *capacity* to refer meaningfully, and hence preclude their very *possibility of meaning*.

# 14

## Horizons

*Nur in der Beschränkung zeigt sich die Möglichkeit.*

– The author’s liberal paraphrase,  
from Goethe’s sonnet  
“Natur und Kunst”<sup>203</sup>

Our imprisonment consists exclusively of the fact that we are not aware of being in our prison. So we cannot (logically) both be in this condition and know that we are in it, and knowledge of our condition is instantly delivering, like the cure for a disease which consists only in not having the cure. But deliverance is complicated by the extraordinary difficulty of explaining to the prisoners, in terms intelligible to them, that prisoners are what they are. For the conditions which make self-understanding possible are incompatible with the conditions they are in, and he who speaks of imprisonment to prisoners must be regarded by them as a madman in his raving. For the bonds which hold them captive are the boundaries of the understanding, and how are we to bring the boundaries within themselves to make them understood? The limits of understanding are not part of what is understood.

– Arthur C. Danto (1971/1968, p. ix)

**W**hen walking along a level beach on a clear and calm day, as you look out over the ocean, if you happen to know the radius of the Earth and the height of your eyes above the ground, with a little trigonometry you can

---

<sup>203</sup> Goethe (1969/1800, p. 245). Very liberal English translation of this liberal paraphrase: “Only within a system of limitation does possibility show itself.” The original lines of the poem are:

*In der Beschränkung zeigt sich erst der Meister,  
Und das Gesetz nur kann uns Freiheit geben.*

calculate how far away the horizon is. The horizon is the limit beyond which you cannot see from your location. If you're watching a distant boat as it moves toward the horizon, it will be seen gradually to disappear. In this sense, a horizon defines, from your present standpoint, how far you can see. It defines a *limit* based on your present position.

But a horizon has a second property: If you move toward the horizon, you will, weather and waves permitting, see more than you did before, as the limit of what you can see correspondingly moves away from you. Or, if you move vertically, you will also see more than you did before; your horizon enlarges. A horizon has this second property of *unboundedness*. On the one hand, a horizon defines a *limit*, but, on the other, the limit it defines has no set *boundaries*.

The English word 'horizon' and these two properties that define its customary geographical or physical meaning have their distant Greek origins in a blending of *ὄρος* (*horos*) meaning "boundary or limit," *ὀρίζειν* (*horizein*) meaning "to bound or limit," and *ὀρίζων* (*horizōn*) and *κύκλος* (*kyklos*) meaning a "bounding circle." Welch mathematician Robert Recorde, the originator of the equal sign ' = ', expressed the following meaning that the word had come to have by 1556: "The Horizonte is a cyrcle whiche parteth that parte of the worlde that wee see, from that whiche wee see not" (Recorde, 2012/1556).

This "circle" that defines a limit with no set boundaries is of special interest to the metalogic of reference, , as the sub-title of this book would suggest. The meaning of 'horizon' that I shall propose has some distant connections in the past that are worth our trouble to review.<sup>204</sup>

In 1683, Leibniz sent a treatise entitled *De l'horizon de la doctrine humaine* to the Académie des Sciences in Paris. In it, he used the concept of horizon to develop the following thesis: The human species depends upon language to express its thoughts; expressions of thought in language are always finite; therefore the sum total of the truths that make up human knowledge and which mankind is *capable* of expressing in language must be finitely limited. Although such limitation is inescapable, the growth of knowledge over time enlarges its "horizon," but this horizon will always be finitely limited.

Using combinatorial algebra, Leibniz sought to calculate the number of combinations of possible words, and the number of books physically required to contain them. He then used this result to estimate the total library that

---

<sup>204</sup> For a discussion of the use of the concept of horizon by philosophers other than those considered in this chapter, see Appendix I, "The Concept of Horizon in the Work of Other Philosophers."

would plausibly be required in order to express all truths that can be known and expressed in language by mankind. Due to the necessary finiteness of the expression of human thought in language, the conclusion Leibniz drew is that a time will be reached—assuming the human species endures long enough<sup>205</sup>—when no truth can be expressed that had not already been expressed.

Let us content ourselves with having encountered a species of horizon, which limits the human doctrine [consisting for Leibniz of all truths that human beings are capable of expressing in language], and of having elevated our mind to reflections which make it recognize in some fashion those limits which nature has given it. (Leibniz, 1991/1693-1715, p. 53)

Leibniz took this one step further, and a significant step: No matter what—however large—finite number is calculated to represent the possible number of expressible human truths, a larger number can easily be constructed (for instance, just by adding 1 to it). In Leibniz's companion short treatise, *Sur la calculabilité du nombre de toutes les connaissances possibles* [*On the Calculability of the Number of All Possible Truths*], written c. 1693-1694, he expressed this conclusion in the following way:

[L]iterally a day will come, on which one can say no more than what has been said already, *nihil dici, quod non dictum sit prius* [nothing can be said which has not been said before]. For, where one says that which has already been said, or even if one wants to continue to say new things, one will exhaust that which still remains to be said, since this is finite, as we have already shown. *It is a case then of giving a number which is greater than the number of all which can be said or pronounced.* It is this which we set out to do. (Beeley, 2003, p. 101, italics added)

In short, Leibniz's "horizon" that limits the unbounded range of humanly expressible truths formulates, in a sense, an early theorem of conceptual

---

<sup>205</sup> Written in the margin of his treatise, Leibniz wrote, and did not quite succeed in crossing out, this comment: "*peu de probabilité que le genre humain dure toujours*" [little likelihood that the human race will continue forever] (Leibniz, 1991/1693-1715, p. 145).

limitation.<sup>206</sup> The “horizon” Leibniz pointed to possesses, in a certain sense, the two special properties that we’ve mentioned: It identifies a range of values—the number of truths expressible by mankind in language—that is both *finitely limited* and, at least for a very long time to come, *without boundaries*. We may therefore speak, on the one hand, of the *limits* of a horizon, and, on the other hand, of the *boundaries* of a horizon, understood in the sense of the extendability in principle of a horizon’s scope of inclusion (although this extendability may, in Leibniz’s view, itself reach a finite limit—given enough time!). I shall call the latter in-principle extendable type of boundaries ‘*Leibniz boundaries*’ to distinguish them from two other kinds of horizon boundaries that we shall come to. Like the physical horizon that defines what we can see when looking out over the ocean when standing on a beach, the scope of what we can see, or in Leibniz’s meditations what we can think and express, may generally be enlarged, but always that scope remains finite and therefore limited.

Nearly 90 years later, in 1781, Kant published his *Critique of Pure Reason*. In the *Critique*, Kant used the word ‘horizon’ (*Horizont*) in several different ways: in the conventional geographical sense (A760/B788); in the sense that a concept establishes a kind of standpoint from which things may be viewed (A658/B686); in an unusual sense that he intended to refer to genera (“horizons”), species, and subspecies (“smaller horizons”) (A658/B686); and in an epistemological sense, which is of interest to us here.<sup>207</sup>

Like Leibniz, Kant used the concept of horizon to express the two properties we’ve identified, differentiating them by using the two German words ‘*Schranken*’, meaning limits, and ‘*Grenzen*’, meaning boundaries, or borders:

If I imagine the Earth’s surface (as it appears to the senses) as a flat plate [*als einen Teller*], then I cannot know how far it extends. But experience teaches me that, no matter where I go, I always see a space surrounding me in which I can proceed further; and thus I know the limits [*Schranken*] of my

---

<sup>206</sup> Michel Fichant, in his commentary to Leibniz’s *De l’horizon de la doctrine humaine*, makes a similar observation: See Leibniz (1991/1693-1715, p. 144)

<sup>207</sup> In his *Logic*, Kant made still another use of the concept of horizon. There, he used the concept to speak of the “horizon of knowledge,” “by which is to be understood the *commensurateness of the magnitude of all cognitions with the capacities and ends of the subject*” (Kant, 1974/1800, p. 45). He went on to distinguish the “rational horizon” from the “historical horizon.” In his view, the latter has no boundaries since historical cognition is always open-ended; the rational horizon is “fixed.” “[W]e can determine, for example, to what kind of objects mathematical cognition cannot be extended. So also in respect to the philosophical cognition of reason: how far reason may advance *a priori* without any experience” (p. 46).

actual knowledge of the Earth, but not the boundaries [*Grenzen*] of all possible descriptions of the Earth [*aller möglichen Erdbeschreibung*]. But if I have gotten so far as to know that the Earth is a sphere, and that its surface is spherical, I can know *a priori* and determine based on principles, from my knowledge of a small part of this surface—for example of the size of a degree—the diameter and from this the circumference of the Earth; and although I am ignorant with respect to the objects which this surface contains, I am not with respect to their extent, size, and limits [*Schranken*].

The sum of all the possible objects of our knowledge seems to us to be a level surface, with an apparent horizon [*Horizont*]<sup>208</sup>—namely that which forms its whole compass, and which has been termed by us the idea of unconditioned totality. To reach this limit by empirical means is impossible, and all attempts to determine it *a priori* according to a specific principle have been in vain. However, all the questions raised by our pure reason relate to that which lies beyond this horizon [*Horizonte*], or at least lies on its boundary line [*Grenzlinie*]. (A760/B788)<sup>208</sup>

The meaning of ‘horizon’ in the above passage shifts from a geographical analogy to the idea of horizon as circumscribing the totality of possible objects of knowledge. But the shift to this meaning of ‘horizon’ is accompanied by an equally evident but still only intimated *desire to step beyond* what in this study I’ve called the ‘boundaries of possible meaning’. In the long passage quoted above, Kant made use of the distinction between “limits,” on the one hand, and “boundaries” (those which I’ve called ‘Leibniz boundaries’), on the other. But he then indirectly introduced a third *wished-for* property of horizons that we need to pay special attention to. Here is the central paragraph in the *Critique* in which he did this:

[T]here is an advantage that can be made comprehensible, and at the same time of interest, to even the most difficult and most unenthusiastic apprentice of such transcendental inquiry, namely this one: that the understanding which is concerned merely with its empirical application, and which does not ponder the sources of its own knowledge, may very well advance, but one thing it cannot do; it itself cannot

---

<sup>208</sup> Author’s translation from the second edition.



determine the boundaries [*Grenzen*] of its use, and cannot know what may lie inside or outside its total scope; because to do this requires the deep investigations that we have undertaken. But if it cannot distinguish whether certain questions lie within its *horizon* [*Horizonte*] or not, then it will never be sure of its claims and its possessions, but can only count on a variety of shameful rebukes when it oversteps [*überschreitet*] the boundaries [*Grenzen*] of its domain (as is inevitable), and goes astray in *delusion* and *deceptions*. (A238/B297, italics added)<sup>209</sup>

We've seen how Kant made use of the distinction between *Schranken* and *Grenzen*, but in the above passage, as I understand him, Kant came close to using the concept of horizon to delimit and to separate what is possibly meaningful from that which is delusional and deceptive. He made no attempt to demonstrate in a logically compelling way that transgressions of such a horizon really do lead to delusion and deception. Nevertheless, this new and radically different sense of "horizon" to which I think he came to refer, he not only claimed *can* be overstepped or transgressed, but he himself—immediately and without checking himself—began to make this transgressing step. We see this in his choice of words when he wrote, "all the questions raised by our pure reason relate *to that which lies beyond this horizon...*" (italics added), when he endowed noumena with reality, as well as in other central claims in the *Critique*. On the one hand, his extended conception of horizon purports to establish a boundary beyond which lie only delusion and deception, but, on the other hand, Kant succumbed in his *Critique* to what, two years later, he admitted is a *compulsion* to violate those very boundaries. These horizon boundaries are clearly of a fundamentally different sort compared with Leibniz boundaries, and so to distinguish them let us call them '*Kant boundaries*'.

Just two years after the publication of the first edition of the *Critique*, Kant published his *Prolegomena to Any Future Metaphysics*. In the *Prolegomena*, like Leibniz, Kant was interested in much the same phenomenon—for Kant, the observation that human reason is limited, but in a way that has no boundaries—i.e., boundless understood in Leibniz's sense of extendability.

---

<sup>209</sup> Author's translation from the second edition. It is unfortunate that both Norman Kemp Smith and J. M. D. Meiklejohn chose to ignore what I take to be Kant's fundamentally important word choices here: '*Wahn*' and '*Blendwerke*', which Kemp Smith translates as "opinions that are baseless and misleading," and Meiklejohn translates as "fanciful opinions and blinding illusions." On the contrary, *Wahn* and *Blendwerke* mean literally that *delusions* and *deceptions* are involved; "opinions" are not involved.

And yet, the willingness to overstep the other, newer kind of “Kant boundaries” identified in the *Critique* had by the time of the *Prolegomena* become *legitimized* in his mind, and so he gave the following, certainly objectionable, rationalization for transgressing these very different boundaries. He wrote:

[W]ho does not *feel himself compelled* [*fühlt sich nicht notgedrungen*], notwithstanding all interdictions against losing himself in transcendent ideas, to seek rest and contentment beyond all the concepts which he can vindicate by experience...” (Kant & Beck, 1950/1783, §57, italics added)?

He then went on to try to make this “compulsion” palatable and acceptable—this “feeling of being forced [*notgedrungen*]”—first by claiming that, in mathematics and in natural philosophy, human reason has limits (*Schranken*), but not boundaries (*Grenzen*), because “something *indeed lies without it* [*etwas außer ihr liege*], at which it can never arrive” (§57, italics added). Here, for human reason, Leibniz boundaries do not exist because Kant judged that “something indeed lies without it.” But then he went on to claim that we are

... *led* [*geführt haben*], as it were, to the spot where the occupied space (viz., experience) touches the void (that of which we can know nothing, viz., noumena [*Noumenis*])... For in all boundaries [*Grenzen*] there is something positive (e.g., a surface is the boundary of corporeal space, and is therefore itself a space, a line is a space, which is the boundary of the surface, a point the boundary of the line, but yet always a place in space), *whereas limits* [*Schranken*] *contain mere negations*. The limits pointed out ... are not enough after we have discovered that *beyond them there still lies something* (though we can never cognise what it is in itself). (§57, italics added)

These horizon boundaries are no longer Leibniz boundaries, but clearly they are Kant boundaries. Here we find the third property of the Kant’s concept of horizon: It is the property that “leads” one, he said, thanks to an inner “feeling of being forced,” to transgress beyond the boundaries of possible knowledge—and so to embrace that which, by his own judgment, is delusional and a deception.

Kant has, of course, been criticized for the evident and extreme incon-

sistency of making claims about that which transcends the boundaries of possible meaning and knowledge.<sup>210</sup> Roy D. Morrison, for example, has referred to the limits of human knowledge as “the transcendental horizon,” and has had this to say, emphatically and rhetorically, about the widespread human urge to overstep these boundaries:

Within the human mind there is a powerful, perhaps primordial, desire to go beyond the *limits* of human knowledge.... Many, many lay persons and professional thinkers invest significant resources in a lifelong attempt to cross the transcendental horizon of the human situation and live in another world or dimension—and to grasp a sense of meaning and security that are not immediately generated by empirical reality alone....

[T]he human mind, apparently, will go to *any lengths* in order to gratify the need for this particular kind of transcendence. No amount of sophistry, no amount of internal contradiction of methods, and no amount of contradiction of empirical reality seem to function as an effective deterrent. In many cases, the *non-recognition of empirical reality* and the rejection of formal logic are basic methodological instruments in the pursuit of meaning, transcendence—and escape. (Morrison, 1994, p. 352)

Before ending this section relating to Leibniz and Kant, it may contribute to the clarity of my purpose if I add the following disclaimer so as not to be misunderstood by reason of misplaced emphasis. This book is evidently not an expository or critical study of the thought of Leibniz or Kant—or of that of any other philosopher. Like all attempts to root out the intended meanings of dead authors—especially an author like Kant who could most certainly have profited by a demanding course in clear nonfiction writing—it can be a challenging and indeed often futile argumentative task to know what he or she really meant. And so the views I ascribe to other philosophers are intended purely as aids to readers—as stepping stones, as didactic tools, as illustrative sample cases—to facilitate an understanding of the shift of conceptual framework required by the metalogic of reference. Whether, for example, Leibniz or Kant specialists generally agree or not with the interpretations offered here of the works of these philosophers—and I say this without diminishing the value of solid scholarship—is irrelevant to the objectives of this study.

---

<sup>210</sup> Including criticism by the author, e.g., Bartlett (1970, Section 2.1, B).

### 14.1 Metalogical horizons

We have looked at the concept of horizon in terms of its common physical-geographical meaning, and then in terms of the ways the concept has been used by Leibniz and Kant. Two defining properties of the concept of horizon have been distinguished in the views of Leibniz and Kant, and a third property which I've identified with Kant's compulsion—"the feeling of being forced"—to overstep the horizon of possible human knowledge.

We turn now to develop a more rigorous and radically different concept of horizon, one which, while it is characterized by the properties of limits and boundaries in Leibniz's sense, is properly metalogical. The metalogical concept of horizon defines certain limits and *a specific and special boundary which, when attempts are made to overstep it, results in projection*. As we shall see, this metalogical horizon determines not only limitative boundary conditions of possible knowledge, but also the limitative boundary conditions of the very nature of possibility, of meaning and reference, of identifiability, of intelligibility, and of ontology.

### 14.2 Limits of reference and boundaries of possibility and meaning

If something were to lie, in principle, beyond possible experience, it could be neither said nor thought nor asked.

– A. J. Ayer (1952/1936, p. 76)

We have described limits and boundaries in connection with the geographical-physical concept of horizon, and also both of these in connection with the concept of horizon employed by Leibniz and Kant. Both Leibniz and often Kant relied upon the common understanding of the limits of a horizon as circumscribing a finite scope of reference, as well as the notion that a horizon is unbounded in the sense that it has no fixed borders, but is frequently potentially extendable, for example, as the observer changes his or her location, or as a series is continued, or as the human species adds to its knowledge.

And then in Kant we also saw a shift to a different meaning of a horizon's boundaries, so that in his view it became possible to speak of "transgressing" or "trespassing" "outside" a horizon's boundaries. This second sense of a horizon's boundaries will now become a central focus of this study. But where Kant merely claimed without compelling justification that overstepping such boundaries was a "transgression" (which, as we know, he then freely engaged in), we shall instead prove that transgressions of properly metalogical horizons result in real incoherence, in meaninglessness. It is a variety of

meaninglessness which, when asserted or believed, qualifies, as we've noted, as real, non-metaphorical *delusion*. —It is a form of delusion in relation to which we are able to comprehend both pre-analytical putative meaning and post-analytical meaninglessness.

Do these metalogical horizons that concern us possess frontiers or borders—that is, fixed lines of delimitation which, if one attempts to breach them, result in projection? If such horizons have fixed boundary lines of demarcation, in what sense might this be so?

Intuitively, there appear to be four distinguishable properties of a metalogical horizon. (From now on, the concept of horizon will be understood only in its metalogical sense.) One property relates to (a) the factual acknowledgement that a horizon circumscribes a finite known set of objects of reference. As we've seen, this is the property that defines the *limit* of a horizon. A second property relates to (b) the metalogical recognition that such a horizon circumscribes a *possible* set of objects of reference; this is the set of objects to which reference *in principle* can obtain from the standpoint of the reference frame whose horizon is in view. And a third sense relates to (c) the misplaced Kantian belief that there exists a boundary, a frontier, a border, that separates the contents of (b) from “all else.” We see immediately that the attempt to establish reference to such a sharply drawn frontier is in principle impossible. For in attempting to specify such a boundary, it instantaneously, so to speak, by logical necessity, becomes part of the contents of (a), and certainly of (b), and therefore ceases to qualify as a “boundary-in-the-putative-sense-(c).” The concept of a metalogical horizon that possesses a determinate boundary line is therefore without possible meaning. Metalogical boundaries are therefore without specifiable, determinate borders.

And yet horizons do unmistakably have boundaries that become evident when reference becomes projective: In this sense, (d), horizons have boundaries in the “symptom-based” sense that reveals to us that something has gone referentially wrong. To use a close comparison, we recognize that an electrical circuit has a fault when it short-circuits. When we detect a short-circuit, this serves as an unambiguous *symptom* of an electrical malfunction. In much the same way, when we detect a projection, we recognize that reference has fundamentally malfunctioned, and in this sense a boundary of meaningfulness has been made evident to us. There is no boundary in sense (c) above, but in recognizing that an instance of reference is projective we detect a symptomatological boundary in sense (d). We might call our recognition of these metalogical boundaries ‘*reactive*’ since such boundaries become detectable only as a reactive result of our recognition that a projection is in view.

In earlier chapters a general theory of possibility was developed that pro-

vides a metatheoretical framework in terms of which structural/ systemic presuppositions and presuppositions of identification can be described, analyzed, and employed within the context of transcendental self-referential argumentation. When we identify the preconditions of reference that are in force in a given frame of reference, we make explicit what is possible from that standpoint. We have recognized referential consistency as a necessary, intrinsically determined general criterion of meaning; affirmed the logical priority of the bonded pair, reference-and-meaning; and seen that any object of reference, any particular, has ingredient in it, embedded in it, as an integral constituent of its identity, the constitutive structure of whatever reference frame forms the basis for its identifiability.

So understood, any object of reference—and here this means its very identity—is recognized to be a function of the set of parametric constraints that permit its identifiability. That set of constraints functions as a group of structural/systemic presuppositions that must be satisfied in order for reference and associated meaning to obtain. The understanding of reference and of objects of reference that we come to is a systems understanding in terms of referential fields. The field concept as applied in this study, as we saw in {10.2}, provides an essentially functional-relational understanding of the organization, the interrelational ordering structure, of systems of reference. In recognizing the field-based nature of systemic totalities, we come to recognize that objects of reference and the reference frame(s) in terms of which they are identifiable are indissociable from one another.

Associated with a given reference frame is a referential field in terms of which possible objects of reference are identifiable. The identity of an object of reference is then understood as an “instantiation” or “realization” of the referential field of its presupposed reference frame. In this specific framework-relative sense, the referential field systemically/structurally presupposed by any object of reference is “self-enclosed.” That is, the referential field always has finitary limits. It also may have Leibniz boundaries that potentially may be extended indefinitely to accommodate, for example, changes in the observer’s location or perspective. And yet such a referential field, while it has no delimited boundaries and is in this sense unbounded, it does possess “reactive” boundaries of a different kind and so is constrained by its metalogical horizon. An analysis in these terms of projectively self-referentially inconsistent concepts or claims makes evident that, although the field has no boundary lines of demarcation, it does possess boundaries which, if violated, undermine possible reference and meaning. Those metalogical boundaries are “made visible,” so to speak, only when such violations occur.

From this standpoint, how then do we come to detect the “short-circuiting

of meaning” that informs us that a metalogical horizon has been violated? We clearly stand in need of a method to detect projections.

### 14.3 The detection of projections

In {8}, we distinguished three general ways in which presuppositions may be examined: as a function of truth, as a function of meaning, and as a function of identifiability. We saw that in the case of truth-functional referential presuppositions, we lack an effective method that can test whether a purported presupposition really is presupposed by a given statement or group of statements. Similarly, we are in need of a method to enable us to recognize presuppositions of meaning and presuppositions of identifiability. We’ll find it useful to devise tests to meet these needs, and, in the process, we shall be able to develop a method to detect projections.

In {8.5} the notion of “destructive testing” was introduced. Destructive tests are particularly useful in connection with integrated, dynamic systems. Whether a certain organ in human anatomy is indispensable to life, or whether a particular line in a computer program is necessary for the program to function as intended, these can of course be tested destructively: If eliminating that organ or that line of computer code brings about either system’s collapse, its indispensability is unambiguously demonstrated.

Presuppositions serve essential supportive roles in the systemically organized contexts in which we determine the truth-value of referring statements, the meaning of concepts, or the identifiability of certain objects. To say that presuppositions serve “supportive” roles is to direct attention to the fact that, if such “support” is taken away or in some sense “destroyed,” the system in question will undergo a form of collapse. The kinds of collapse that are involved will differ according to the nature of the presupposition and its role in the system in which it is presupposed. In this sense, destructive testing lends itself well to the analysis of presuppositions of whatever variety.

In {8.1}, 13 varieties of presupposition were distinguished. Of these, we noted that five of these have been of particular interest to philosophers: presuppositions of existence and linguistic presuppositions of reference (together comprising referential presuppositions), presuppositions of concepts, structural/systemic presuppositions, and presuppositions of identification (which we regard as forming a sub-class of structural/systemic presuppositions). We then identified the three above-mentioned general ways in which presuppositions may function: as a function of truth, as a function of meaning, or as a function of identifiability.

It is a straightforward matter to devise destructive tests of these ways in

which presuppositions may function:

For *referential presuppositions of truth and falsity*, we recall that if B presupposes A—i.e.,  $B \wp A$ —then the presupposition relation  $\wp$  holds only when (1) if B is true, A is true, (2) if A is false or is not satisfied, B is neither true nor false (and if B is a question, command, etc., it is inappropriate, senseless, etc.), and (3) if B is false, A must be true or be satisfied. To test whether a referential presupposition relation holds between B and A, we simply deny that A holds or is satisfied. If doing this has the logical effect that B can have no referent, then this destructive test has succeeded. B presupposes that A holds, otherwise B can have no referent, and hence cannot be either true or false. If B is “The back door of this house is locked” and A is the presupposition that “This house has a back door,” then if A does not hold, B is neither true nor false; it may be inappropriate, odd, laughable, etc., but it is not true, and not false.

For *presuppositions of concepts*—for example, that the concept of hawk has ingredient in its meaning the concept of bird—again, we test the presupposition relation by taking away the allegedly presupposed concept and then note the effect. A hawk that is not a bird is not, given the meaning of the concept of hawk, what is meant by that concept. Without the concept that is presupposed, the presupposing concept’s meaning collapses.

In the case of *systemic/structural presuppositions and presuppositions of identification*, our interest in this study specifically concerns the forms these take in transcendental terms, i.e., when our interest is focused upon their metalogical forms—upon what must be granted in order for a system or structure or an object’s identification to be possible. In {8.1.13, 8.6}, we also distinguished two varieties of presuppositions of identification, those that are object-related, having to do with identification criteria that objects of reference must satisfy in order to be identifiable in principle, and system-related, having to do with the frame(s) of reference in terms of which identification can obtain. Let us take these one at a time:

To test whether an alleged presupposition of identification is necessary in order for it to be possible for a certain object or set of objects of reference to be identified, we recall that presuppositions of identification are understood in the following way: If  $S$  is an identifying description of an object  $o$ , and  $P$  is a condition that must be met in order for  $S$  to obtain or succeed in identifying  $o$ , then ‘ $S$  presupposes  $P$ ’ means that for  $S$  to identify  $o$ ,  $P$  is a necessary condition of  $o$ ’s identifiability. Such a  $P$  is what is meant by a presupposition of identification.

To test whether  $P$  is a necessary condition of  $o$ ’s identifiability, we assume that condition  $P$  does not hold or is not satisfied: For example, con-



sider an identifying description of a certain person named ‘O’—e.g., the description given by “O is the man who has a birthmark on his left hip, is currently 27 years old, has Social Security number \_\_\_\_, lives at the address \_\_\_\_.” Suppose further that we wish to know whether this identifying description presupposes the following partial list of conditions—several *Ps* in this case—of O’s identifiability: These conditions include, for example, the ability to distinguish human males from human females, the ability to detect a birthmark on a person’s left hip, access to Social Security records including ages of individuals and their addresses, etc. To test whether such *Ps* are indispensable to the possibility of identifying O, we assume these conditions are not met, and immediately the identifiability of *O-as-described* ceases to be possible. There is an inescapable presuppositional relation between O’s identifying description and the conditions that would need to be satisfied in order to make O’s identification possible.

Similarly, in order to test whether a system-related condition of identifiability holds, we examine the frame of reference in terms of which identification is presumed to obtain. For example, suppose we consider the GPS coordinates of a physical location, and wish to know what systemic conditions must be granted in order for that location to be identifiable. Assuming, for simplicity, that we agree to limit the scope of this example to the global positioning system in terms of which GPS coordinates have reference and meaning, a few of the relevant presupposed systemic conditions include the following: the ability to take and compare very precise measurements of the time that electromagnetic transmissions require between a relevant subset of global positioning satellites and ground locations, the ability to integrate this time measurement information in a computer program that incorporates the principles of general relativity, and the ability to display the results of the program’s calculation in terms intelligible to individuals or machines that accordingly identify the specified physical location.

This exercise in listing some of the presupposed conditions of identifiability that pertain to an overall, interdependent system such as the global positioning system enables us to recognize in less abstract, applicable terms how such presuppositions of identification are structural/systemic rather than truth-functional in nature. They are appropriately termed metalogical in the sense in which they comprise preconditions that must logically hold in order for an object of reference *possibly* to have a certain identity (in the above example, the GPS-designated identity of the specified physical location). They render *possible* that which, without them, would collapse—much like the dynamic system of stresses and strains of a suspension bridge is subject to being tested by the strength of the cables that support and prevent the bridge from falling.

In the preceding example, if *P* represents the role of general relativity in the GPS system, *P* serves as a metalogical condition of the *identifiability* of the GPS-designated position *S*, and if there is such identifying reference to *S*, then *P* must *necessarily* hold or be satisfied; if condition *P* does not hold, such identifying reference to *S* is, from the standpoint of the GPS system, theoretically *impossible*. If *P* is denied, if the condition that *P* represents does not hold, then *S* cannot have *possible* reference and *S* cannot have *possible* meaning; the two form a bonded pair.

A projection in this example results if it were to be claimed that the *GPS-determined location* could be identified in a manner *autonomous* of the GPS system, while yet claiming at the same time that the conditions required to do this are not satisfied. Such a contrived claim should not be misunderstood as the very different and unobjectionable claim that the same physical location could be identified by other, non-GPS means (say, by using maps, landmarks, traditional latitude and longitude measurements, etc.). A conceptually malfunctioning claim here that is projective is a claim that, on the one hand, putatively depends for its *possible* meaning and reference upon a certain frame of reference (the GPS system itself), and yet involves both (1) a denial of one or more metalogical presuppositions of identification/identifiability *necessary* to that frame of reference, and (2) an assertion of the *autonomy* or *separability* from that frame of reference of a putative identification. To do this in a way that combines both (1) and (2) is self-undermining on a metalogical level and is clearly projective.

In considering an example like this, as has been underscored in previous chapters, the reader should again take note of the central role in the preceding paragraphs of the modal concepts of possibility and necessity; a modal level of analysis characterizes the level of analysis appropriate to the metalogic of reference.

The tests I have outlined in order to enable us to determine the validity of referential presuppositions of truth and falsity, or of presuppositions of concepts, or of systemic/structural presuppositions and presuppositions of identification, all have the same structure: They all proceed destructively, by removing one or more presuppositional “supports” to determine whether a given statement, proposition, concept, structure, or system will then fail.

When metalogical presuppositions of identification/identifiability and metalogical systemic/structural presuppositions of identification are destructively removed or denied, as they are in projections, horizons of possibility and meaning are made visible to us. Destructive tests for the presence of such metalogical presuppositions reactively disclose the metalogical boundaries with which we are concerned, boundaries that make evident to us what I have

called ‘*horizons*’. In this way we become aware of what I earlier have called the ‘thresholds of meaningfulness’, whose mapping is one of the main tasks of the metalogic of reference.

#### 14.4 Heuristics and the detection of horizons

[M]any *heuristic* [*heuristische*] methods of thinking lie perhaps still concealed in the experience-use of our understanding and of reason, which methods, if we understood to draw them carefully from that experience, might enrich philosophy, even in abstract cogitation....

– Immanuel Kant (1798-1799, vol. 1, p. 387)

A process that *may* solve a given problem, but offers no guarantees of doing so, is called a *heuristic* for that problem. This lack of a guarantee is not an unmixed evil.

– Newell, Simon, and Shaw (1957, p. 220)

Presuppositions tests like those described in the previous section evidently function as *heuristic tools of analysis*. They help to establish metatheoretical rules to guide subsequent analyses. It is important that we understand the rationale for using a heuristic approach.

The word ‘heuristic’ is associated with the well-known legend according to which the famous Greek mathematician, physicist, and inventor Archimedes cried out “Eureka!” (in Greek: εὕρηκα [*heúrēka*]) when he found the solution to a problem he had been assigned: to determine if King Hiero II of Syracuse had been cheated by goldsmiths who may have diluted with cheaper silver the gold which the King had given them with which to make him a crown. The Greek “*heúrēka!*” that Archimedes exclaimed means “I have found [the solution]!”

In the many centuries since Archimedes, ‘heuristic’ has come to refer to a method, rule, or process that can serve the objectives of problem-solving, but one which, as Newell, Simon, and Shaw observed, does not *guarantee* a solution. A heuristic method is less formally structured than an effective algorithm that moves from an initial problem state to a final valid solution without any uncertainty that a solution will be found. In general terms, a heuristic method is a set of operations which, when applied to a problem suited to that method, and provided that the method’s application is successful, will help one, if not

to solve it, at least to come closer to a solution. Heuristic methods are not necessarily to be looked down upon, for they can be invaluable when precise and assured methods are too cumbersome, slow, often simply not yet available, or, perhaps, impossible. When Newell, Simon, and Shaw remarked that the lack of a guaranteed solution “is not an unmixed evil,” such potential advantages of a heuristic approach were what they had in mind.

When Einstein received the Nobel Prize in physics in 1921, it was for a 17-page paper that he published in 1905, “On a heuristic [*heuristischen*] point of view concerning the production and transformation of light” (Einstein, 1905).<sup>211</sup> A heuristic approach is not necessarily to be scorned.

With respect to the metalogic of reference, I have chosen to develop it in terms of heuristic tools of analysis for a number of reasons, foremost among them are these: As Part I of this book sought to make clear, philosophers generally tend to be a contentious lot; the arguments they propound are formulated in ways designed to resist criticism voiced by their competitors, and such arguments are, as a result, often stated with less than desirable precision—given that obscurity and ambiguity can serve, whether consciously or unconsciously, the practical purpose of defending against damaging criticism. In a discipline in which contention and position-taking often dominate, an adaptable heuristic approach, which can more easily accommodate and be applied to the multiple frames of reference of philosophers, does not suffer from the rigidity of a proposed strict algorithmic formulation whose limitations, those of any rigorously focused algorithm, can easily be targeted by criticism and dismissal.

Second, many of the problems to which our metalogical rules and guidelines of analysis can appropriately be applied are problems that require a form of analysis that proceeds with a open and flexible problem-solving attitude; such problems benefit from reflective study from “a heuristic point of view,” to use Einstein’s phrase, a point of view that seeks “to *find*” (in Greek: “*heuriskein*”) solutions, rather than to *propound* them.<sup>212</sup> As we shall see in Part

---

<sup>211</sup> Einstein’s paper made use of the term ‘heuristic’ only in its title, but the content itself of his paper is heuristic in nature: It formulates a series of (very prescient) hypotheses, which, if borne out by subsequent experiments, were intended to *find* a better understanding of the photoelectric effect—which indeed he succeeded in doing.

<sup>212</sup> Heuristic is often described the way mathematician George Pólya did many years ago: “The aim of heuristic is to study the methods and rules of discovery and invention” (1971/1945, p. 112; see also 1941, 1950, 1954, 1962-65). I do not accept this conception due to the involvement of projections in the naive notions of “discovery” and “invention,” as we shall see in Part III. For this reason, heuristic and heuristic methods are best understood as ways that enable us “to *find*” a solution to a problem—that is, ways to reach a *recognition*, *acknowledgment*, or *identification* of its solution.

III, such problems are of such considerable conceptual diversity to a degree that it may be difficult, and perhaps even theoretically impossible, to devise strict, effective algorithmic methods to deal with them—methods that “guarantee a solution.”

A third and last motivation that may be mentioned at this point in favor of a heuristic approach is this: It is an important question to ask to what extent the approach of the metalogic of reference is itself subject to adequate—I will not say “complete”—deductive formalization. Related to this is the question whether the detection of projections could in principle, in any significant sense, be automated. We shall consider such questions in a later part of this book (see Supplement).

### 14.5 Reflections on horizons

In this chapter, we’ve come to understand the concept of horizon in several ways—as defining a finite limit based on one’s present position or point of view; as possessing the property of unboundedness which relates to the potentially indefinite extendability of the foregoing limit; as a scope of perception, of thought, or of experience generally which may encourage the human desire to step beyond the boundaries of possible meaning—or which may promote the “feeling” that one is somehow “forced” like Kant to make such a trespassing step; and as defining a specific and special boundary which, when attempts are made to overstep it, results in projection. When the latter occurs, such transgressions of *metalogical horizons* result in genuine incoherence, in meaninglessness, in real, non-metaphorical delusion.

We’ve recognized four distinguishable properties of a metalogical horizon. They are the properties that define: (a) the *limit* of the finite set of objects of reference a horizon includes, (b) the *modal limit* of possible objects of reference it permits, (c) a horizon’s *delusion-inducing* character that encourages a misplaced belief that there exists some sort of border that separates the possible objects of reference from “all else,” and, finally, (d) the *reactive* or (*symptomatological*) character of metalogical horizons whose boundaries become detectable only as result of the occurrence of projections.

We saw that associated with a given reference frame is a referential field in terms of which possible objects of reference are identifiable, a field in terms of which an object of reference is then understood as an instantiation or realization of that field. While such a field has no delimited boundaries and is in that sense unbounded, it exhibits a horizon’s reactive boundaries that become evident when violations of its horizon occur.

Destructive tests were then developed to enable us to determine when

three kinds of presupposition relations do or do not hold, that is, presuppositions of truth and falsity, presuppositions of concepts, and transcendental systemic/structural presuppositions, including presuppositions of identification. We found that when metalogical presuppositions of identification/ identifiability and metalogical systemic/structural presuppositions of identification fail to be satisfied, metalogical horizons of possibility and meaning are made evident to us. Such destructive tests for the presence of these metalogical presuppositions make evident what in this study I've called '*horizons*'. They make evident certain inescapable limitative *thresholds of meaningfulness*, which, when attempts are made to exceed them, result in *metalogical self-referential inconsistency*.

Finally, we have summarized some of the major advantages of a heuristic, non-algorithmic approach to detecting metalogical horizons.

As suggested in {13.5}, implementing the proposed heuristic guidelines that define the metalogic of reference will ideally lead to what appropriately can be called a '*critique of impure reason*', for in providing us with the ability to recognize preconditions of possible reference, such a critique guides us to a body of results in the sense of Kant's "*negative science*." Such heuristic tools give us a way to determine the most fundamental regulative constraints upon valid reference, and in doing this, they negatively define the domain of possible meaningfulness by allowing us to recognize the horizons which, if one attempts to exceed them, result in meaninglessness and delusion. In Part III, as we apply the results we will have reached to major individual problems that have occupied philosophers, we shall have the opportunity to add progressively to our group of heuristic tools of analysis.

...

Not by any concerted act of imagination can we trespass beyond the boundaries of what for us is imaginable.<sup>213</sup> —This is a tight tautology, within which we realize all the intellectual liberty that is possible for us. The internal limitations of human understanding disclose themselves in several distinct ways: In our practical dealings with the world, we are subject to *neurological* limitations and to limitations of *language* and *idea*. And in our conceptual efforts, we are constrained by *epistemological* boundaries.

The picture of the human condition suggested by these limitative factors is one of a finite organism whose neurology is responsive to a range of possible stimuli, whose conceptual vocabulary permits a certain breadth of theo-

---

<sup>213</sup> The last few paragraphs of this section are based in part on Bartlett (1992b), with substantive changes.

retical representation, whose natural and abstract languages allow for a scope of expression and demonstration, and whose extent of knowledge is determined by conditions and limits described by epistemology. This is a picture of a creature who inhabits a specifically human universe of meaning, one which Kant felt forced to believe is but a fragment of a more inclusive reality, from contact with which our practical and theoretical limitations eternally bar us: a more inclusive reality which he called ‘noumenal’.

Appealing though this picture may be to poetic fantasy, it is a grossly distorted, delusional one: It misconstrues the compass and the kind of internal limitation that forms our subject here. This delusional view, which situates human reality within a more comprehensive framework, seeks to *export* and yet *presupposes* the very concepts, language, and neurology that define the human perspective. In this step of attempted exportation, we run headlong into the invisible constraints of metalogical horizons, from which “escape” is not only impossible but, on reflection, also is unthinkable and self-undermining on a level of possible meaning. The existence of these constraints is theoretically determined and does not depend upon the contingent biological, conceptual, or linguistic abilities of a particular organism: In attempting to refer beyond the reality made possible by our neurology, concepts, and language, we attempt, in essence, to refer beyond the reach of our referring capacities. We seek to do the impossible—not the impossible in practice—but *the impossible in principle*.

The metalogical horizons this chapter has described, the “boundaries” of our understanding, are very peculiar boundaries, unlike the boundaries that delimit a parcel of land, or the walls that enclose a box. They much more closely resemble the self-limiting and yet unbounded character of a continuum that has no “outside,” such as is formed by a topologically recurved surface or volume. A close analogy is a topologically closed relativistic model of the physical universe, unbounded yet finite. In such a model, no matter where one goes, no matter how far, there is no way “out.” For the very notion of an “outside” is *part* of the universe of meaning whose internal limitations we should by now begin to appreciate. These “limitations” are of a special, epistemological variety; here the ordinary meaning of the concept of limitation has undergone a radical change.

If we cannot reasonably assert (or deny) that there is an “outside,” lying “beyond the reach” of the powers of our neurological structure, concepts, and languages, then it cannot, in principle, make sense to say that we are *constrained* by these internal limitations. Instead, at the beginning of this chapter, as I paraphrased a line from Goethe’s poem, “*Natur und Kunst*,” it is only within a system of limitation that that which we understand by the possible

and by the meaningful *can be* manifested. Beyond the horizons that bound the universe of the possible and the meaningful lies only unintelligibility, incoherence, and meaninglessness.



# 15

## De-projection

The love of truth is not incongruous with a passion for correcting the erring.

– Gilbert Ryle (1959/1946, p. 329)

Central to the metalogic of reference is what we shall call the ‘*method of de-projection*’. The formal structure of this method expresses a metatheoretical truth established as a function of the principles that follow necessarily, in a self-validating manner, from the nature of framework relativity. This result will provide us with a fundamental theorem upon which our later analyses will be based.

Let us begin by calling ‘*de-projection*’ the method that has the following three objectives:

(1) to *detect* projections when they are present in theories, positions, and concepts;

(2) to *eliminate* projections from such theories, positions, and concepts which claim to be rational—which, in other words, accept the avoidance of self-referential incoherence as an essential methodological value; and

(3) to *re-formulate*, with what we may call ‘*respectful sympathy*’, those projective theories, positions, and concepts in such a way—when this is possible—that seeks to express their pre-analytical intended “putative meaning” so as to avoid projection; this re-formulation seeks, when it can be done, to express the sense which, in reflective analysis, we take, or sometimes must imagine, those theories, positions, and concepts to have in order for them to possess possible meaning.

The reader will note that the purposes of the method of de-projection are stated in terms of its application to theories, positions, and concepts, rather than in terms of its application to the assertions, beliefs, or claims to knowledge expressed by individual human or other agents. We do this to re-emphasize that the method is a heuristic tool designed primarily to be applied to *conceptual structures* rather than to pragmatical uses to which human or other agents put those structures.

Part I of this book explained in some detail the rationale for this choice of analytical focus. Applications of de-projection in a pragmatical context—in the context of the utterances, affirmations and denials, and advocating of claims by individual agents—are certainly possible. It is a level of analysis which the author has frequently engaged in, and which he has alternately found interesting, frustrating, and sometimes pointless, as a teacher of his students and in exchanges with professional colleagues. As long as philosophy does not possess and adhere to a unitary methodology and set of standards of demonstration, and as long as its practitioners possess a general psychological constitution characterized by personal vested interest in their preferred beliefs and in position-taking, applications of any method—even a self-validating methodology as is found in the metalogic of reference—are unlikely to bring about incremental progress in the discipline.

Given the challenges and shortcomings of pragmatical applications, the objectives of de-projection are defined then in terms of the method's application to conceptual structures such as theories, positions, and concepts, rather than to interpersonal rhetorical argumentation.

The first two objectives of de-projection identified above are self-explanatory. The third requires clarification: When a theory, position, or even an individual concept entails projection, there is a need to understand clearly what is meant by “respectful sympathy” for its intended, putative meaning prior to analysis.

To take an example: When, as we saw in the preceding chapter, Kant claimed that “even the most difficult and most unenthusiastic apprentice of ... transcendental inquiry” tends to transgress the boundaries of possible knowledge, and thereby “goes astray in *delusion* and *deceptions*,” he implicitly recognized that such an apprentice *had something in mind*. Despite the fact that an individual is delusional, despite the fact that he or she is so deeply mistaken as to embrace deception, still that person apparently has something in mind, however erroneous or fundamentally incoherent it may be.

To consider such projective delusions in a respectful, sympathetic light means that we make an effort to untangle them, and, when possible—which it not always is—we attempt to substitute for them re-formulations that elimi-

nate and avoid the projective errors they involve. We appropriately call this procedure ‘*de-projection*’: Through its application, projectively self-undermining delusions are thereby “*de-projected*.”

In {11} and {13} the concept of projection was defined in both a formalized manner and discussed informally in some detail. We recall that a pair of conditions must be satisfied in order for the referential short-circuiting that constitutes a projection to occur: First, a particular object of reference is putatively considered as though it were dislocated from the frame of reference that must be presupposed in order for that object to be identifiable. The object of reference is regarded, in other words, as though its essential relativity to the reference frame required for its possibility is *severed*. It is this purported “dislocation” or “severing” of object of reference from metalogically presupposed reference frame that meets the first condition that defines a projective reference. Second, there must be an implicit or explicit claim that the object of reference possesses an *autonomy* with respect to its presupposed reference frame. This second condition, the satisfaction of which takes the form of a claim that the object of reference possesses an autonomy from the identification system(s) to which its possibility is relative, when conjoined with the first condition, results in the special variety of metalogical self-referential inconsistency that I’ve called ‘projection’. When reference becomes projective, reference is putatively made to an object of reference while the grounds for that object’s possible identity and identifiability are ruled out. In short, a projection purports to sever the relation of relativity of an object of reference to the appropriate identification framework(s) in terms of which it can, in principle, be identified.

In the previous chapter, we encountered several properties of constraint that apply to metalogical horizons. These are properties that express: the *limit* of the finite set of objects included in a given horizon; the scope of possible objects to which reference can be made from the standpoint of the horizon—i.e., the horizon’s *modal limit*; the *delusion-inducing* character of a horizon, which promotes the delusional belief that “beyond” the horizon there lie objects the existence of which one may in some sense feel a need to invest belief in; and the *reactive* property of a metalogical horizon, which refers to the manner in which its boundaries become evident only as a consequence of the occurrence of projective attempts to exceed, to transgress, and so to violate what is possible from the horizon’s presupposed reference frame.

Once the three objectives that define de-projection have been realized, we shall find that the delusion-inducing character of horizons is no longer possible and can no longer present a problem. Identifying reference to objects, the very identity of such objects, the parameters of constraint of a reference

frame, possible meaning, and intelligibility are all of them subjected to and governed by the de-projective understanding that comes with a full recognition of framework relativity. By applying the method of de-projection, Kant's "feeling of being forced" beyond the boundaries of possible meaning will, so to speak, be neutralized or silenced. In the process, as we shall see later, a large number of philosophical problems find their strongly compelling solutions.

De-projection, in other words, is a method that forces one, on pain of metalogical self-referential inconsistency, to accept a framework's metalogical horizon—without struggling in a futile and self-undermining attempt to transcend it. Where a projection would *dislocate* purported objects of reference from the reference frame(s) necessary for the possibility of such reference—and by doing so putatively sever contextually relative ties to an underlying reference frame—de-projection provides a method that reintegrates or restores objects of reference within their metalogically presupposed frames of reference, affirming their inescapable framework relativity.

In preceding chapters, I have tried to give the reader some idea of the scope and interests of the metalogic of reference. The route we have taken has led from an initial interest in the approach to conceptual analysis proposed by transcendental philosophy, to an understanding of the concept of possibility in terms of metalogical preconditions of valid reference. From this standpoint, we have sought in particular to understand the metalogic of reference in terms of those ways in which reference fails by undermining its possibility, and in the process, undermining the possibility of meaning, coherence, and intelligibility.

To identify, avoid, and eliminate the variety of invalid reference that I've called 'projective', the method of de-projection rests on the following conclusions we have reached:

- (i) It is impossible to refer meaningfully when the preconditions of such reference are denied.
- (ii) A reference which entails a denial of its preconditions is metalogically self-referentially inconsistent and is to be eliminated and avoided by any coherent, rational approach.
- (iii) Any universally applicable reflection on the metalogical constraints upon reference must accept (i) and (ii) if a coherent metalogic of reference is to be possible.

As the reader may verify, each of these propositions immediately self-validates if denied.

In {5}, {7}, and {10}, we saw that frames of reference function as systems of coordination that provide the basis for the identifiability of a range of possible objects of reference, and thereby form the foundation for our ability to know and to communicate knowledge of any class of objects of reference. In {13}, the concept of projection was defined as putative reference that conflicts with the system of constraints of its presupposed coordinate framework. Understood in this way, we saw that a projection comprises a putative coordination that is invalid because such purported reference conflicts with what must be granted in order for it to be possible. When reflective metalogical analysis shows that this is the case, the putative reference is shown to be devoid of sense; its intended meaning “short-circuits,” exhibiting a reactive symptom that a metalogical horizon boundary has been transgressed.

Within this context, the method of de-projection is designed, then, to make it possible to clarify and, when possible, to restore a given projective reference so that its putative meaning complies with the regulatory structure of its presupposed system of identification.

To accomplish this end, de-projection assumes the character of a *transcendental heuristic* that re-integrates an object of reference within its conditioning context, affirming the essential relatedness—the “intimate tie” discussed in previous chapters—of objects of reference to the frameworks in terms of which they are identifiable.

### 15.1 The heuristic stages of de-projection

We may list four stages that are involved in the de-projective analysis of a theory, position, or concept: *descriptive*, *diagnostic*, *eliminative*, and *corrective*. Let us look at each of these in turn. (For conciseness, in what follows in this section, whenever the term ‘theory’ is used, the phrase ‘theory, position, or concept’ will be assumed.)

When a theory is subjected to de-projective analysis, our first concern is to render explicit the parameters of constraint that determine the range of possible reference from the standpoint of the reference frame it presupposes. There is at present no algorithm to accomplish this first stage of *descriptive* analysis: We need to consider the theory and examine how it functions as a system of coordination, a system that typically establishes implicit or explicit coordinative relations among (i) putative objects of reference, (ii) often a formal or other system that provides a contextual background, perhaps (iii) reference to a time, (iv) position, and/or (v) relationship to a one or more

observers or recording instruments of detection, measurement, calculation, etc. In the descriptive stage of de-projection, we pay special attention to the manner in which the theory serves to enable the identifiability of objects within its scope of reference, and we proceed to describe the referential constraints governing the given framework of reference.

As we shall see in Part III of this study when we come to individual metalogical analyses, the descriptive stage of de-projective analysis requires on the part of the analyst perceptive reflective skills and the ability to recognize and identify a theory's fundamental metalogical presuppositional structure. There is, at least at present, an unmistakable human element upon which the effectiveness of all four heuristic stages of de-projection relies. Later, we shall have occasion to identify and discuss the essential cognitive skills that are involved (see {17.3 and 30.8}, Supplement §10, and Appendix II).

In the *diagnostic* stage of de-projection, the results of the first stage provide us with a basis to determine whether the purported objects of reference to which the theory intends to refer can, in principle—given the parameters of constraint of that theory's presupposed reference frame—comprise possible objects of reference. In this stage, care and methodological strictness are indispensable, especially when a given theory has become so second-nature, habitual, and conventionally accepted that it may require a significant degree of intellectual openness and initiative to consider it in a critical light.

The diagnostic stage of de-projection incorporates the approach described in the previous chapter that permits us both to perform the destructive tests needed to identify a dynamic system's structural/systemic presuppositions and presuppositions of identification, and to detect projections when these are present. We recall that destructive testing is particularly useful in connection with integrated, dynamic systems. Individual theories tend to comprise, or to be linked with, systems that form dynamic totalities; as reference systems they possess a constitutive structure that enables characteristic forms of reference, while, at the same time, often also supplying the potential for projective reference.

The diagnostic stage of de-projection seeks to determine whether such a system's characteristic intended forms of reference conflict with its metalogical presuppositions. The specific intent is, of course, to make explicit, in cases of such conflict, that projections are involved. When a projection is found, the assertion of framework autonomy involved in the projection is formulated explicitly. Finally, the opposition of the projective character of the assertion of framework autonomy is verified to result in a self-undermining and meaningless formulation.

We then proceed to the *eliminative* stage of de-projection only when the

preceding diagnostic stage has established that the theory under analysis does in fact involve projective reference. When we have made the decision that this is the case, it is tempting to wish to eliminate a projection by affirming its negation. But we recall that to do this is fallacious: In {6} and {11} we reached the result that *rejecting* a projection does not authorize an assertion of its negation. A theory's projective claims—claims that are made which initially are believed to be meaningful, and which are therefore, prior to analysis, believed to be either true or false—are such that simply negating them *also* results in projection (as expressed in {11.4}:  $p \Vdash \vec{p}$  and  $\sim p \Vdash \vec{p}$ ). We concluded that because projections “short-circuit” on the level of possible meaning, they are rationally unacceptable, and are to be rejected for that reason. But to *reject* such projections *is not to assert their negations*; it is rather to disallow and dismiss the projective nature of the theory under analysis.

The *corrective*—we might also call it the ‘*restorative*’—fourth stage of de-projection has the task of *re-formulating* the theory in question with what I've called ‘respectful sympathy’ in a manner that eliminates and avoids the theory's intent to refer projectively. De-projection is completed when it is possible to reconcile the purposes which a theory's projection(s) would serve, with a suitable re-formulation in keeping with the described constraints on possible reference. This final phase of de-projection involves a correction of such projections, imposing upon them conformity with the preconditions of their possibility. As already noted, it is not always possible to do this, and, even when it can be done, the result will not always be palatable or acceptable to those who have a deeply rooted interest in continuing to endorse and use such projective conceptual structures.

To correct and restore meaning to a projectively meaningless theory is necessarily to re-conceptualize it along non-fallacious lines. To do this involves an unavoidable interpretation of its original set of purposes. When—in a way that simply cannot be separated or disengaged from the theory—those original purposes *insist* that the theory's purported objects of reference “lie beyond the metalogical horizon” of the theory, it will not be possible to re-formulate the theory in a manner that “restores” meaning where meaning is, in principle, not possible. When this situation arises, we're left with no other option than to conclude that the pre-analytical, intended “putative meaning” of the theory is irremediably devoid of meaning, and must, in the interests of rational coherence and intelligibility, be rejected. We are compelled, on pain of metalogical self-referential inconsistency, to conclude that any attempt to “salvage” meaning in such a case is, in principle, futile.

### 15.2 The epistemological neutrality and tautological nature of de-projection

In {4}, the metalogic of reference was characterized as a maximally general metatheory that seeks to avoid the unreflective introduction of extraneous philosophical presumptions. Here, we shall distinguish the goal of such philosophical neutrality from the *epistemological neutrality* that de-projection assures.

When projections are avoided by responsible compliance with the preconditions of reference that define a given framework of reference, *metalogical analyses introduce no supplementary content or assumptions*. As a method of description, diagnosis, and elimination of projections—that is, comprising the first three stages of de-projection, but omitting the fourth interpretive corrective stage—the formal structure of de-projection is *tautologous*: In making explicit the metalogical constraints that a given reference frame presupposes, a point is reached in reflective analysis where it is possible by self-validation to show that a theory's scope of possible reference metalogically entails the structural/systemic presuppositions of that theory or reference.

In this special sense, the method of de-projection is *empty of content in the tautologous sense*; no content is introduced that is not *already ingredient* in the system of referential coordination that is under analysis.

We find a parallel state of affairs in connection with any accurate description—in the sense in which every accurate description may be viewed as tautologously related to that which it describes: The information contained in an accurate description is tautologously contained in and borne out by that to which the description applies.

The term 'tautology' originates from the Greek word '*tautologos*', which means "repeating what has been said." It is a combination of '*tauto*', meaning "the same," and '*logos*', meaning "word" or "reason." The original sense of "equivalence of meaning" found in the Greek '*tautologos*' has carried over into logic and mathematics where it has come to refer to a logical law, a rule or proposition true in all possible cases or in all possible worlds. A tautology in this sense is necessarily true, and can authorize the logical transition from one proposition to another while insuring their equivalence. By legitimating such transitions, tautologies are empty of content; they are purely formal and say nothing about matters of fact; but as significant expressions of logically necessary truths they are not devoid of meaning.

In de-projection, the accurate description of the preconditions of reference of a given theory bears a tautologous relation to that theory's constitutive structure—to the structural/systemic presuppositions that must be satisfied in order for reference to objects in the given framework to be possible. The



“tautological equivalence” here is metalogical, expressing a metalogically necessary relation of entailment (see {11.5}) between the set of preconditions of reference of a system of reference examined in the first three stages of de-projection, and the referential framework itself.

This tautological equivalence is fundamental to the metalogic of reference. It may be expressed in the form of the *theorem of de-projection* which follows from a recognition of this equivalence and expresses the nature of framework relativity as developed in this study. (It is, to be more precise, a *metatheorem*—that is, a theorem about the metatheory we are developing, i.e., about the metalogic of reference itself.) The following outlines an informally stated, condensed proof of the theorem:

- (i) Let  $T$  be an arbitrarily chosen, rationally committed theory that permits identifying reference to a class of particulars.
- (ii) Let  $M$  be a description of the set of structural/systemic presuppositions that are necessary for the possibility of such reference from the standpoint of  $T$ . (Stage 1 of de-projection)
- (iii) Let  $s$  stand for a claim made from the standpoint of  $T$  that putatively refers to purported objects of reference such that  $s$  entails the denial of one or more structural/systemic presuppositions in  $M$ : i.e.,  $s$  is projective. (Stage 2)
- (iv) As a rational theory,  $T$  metalogically entails the rejection of  $s$ . (Stage 3)
- (v) Taken together, (ii)–(iv) affirm a tautological equivalence relation between  $T$  and  $M$ : This is the form of tautological equivalence that we find in the relation between an accurate metalogical description of the structural/systemic presuppositions of a given theory and the theory itself. The theory entails that set of metalogical presuppositions, and those presuppositions form the preconditions of reference of the theory. To hold either while denying the other is metalogically self-referentially inconsistent. It is important to bear in mind that this relation of tautological equivalence is not reducible to a relation between truth-functional statements or propositions, but rather expresses the structural/systemic functionally interrelated nature of a dynamic referential system as described earlier in {8.3 and 9.2}.

In short, the first three stages of de-projection, taken together, exhibit a formal structure that is reflexive and self-validating.

An application of the first three stages of de-projection leads to an affirmation of tautological equivalence between a given theory's framework of reference and its metalogical presuppositions. The theorem of de-projection informs us that the formal structure of the method of de-projection is tautological, imports no information not already ingredient in the theory under analysis, yet the tautological nature of de-projection is not trivial in the way that mere truisms or redundancies are. De-projective analysis leads to an explicit description of the metalogical presuppositional structure of the theory under analysis.<sup>214</sup>

Because de-projection is empty of content in the tautologous sense, it can authorize the error-free transition from the metalogical description provided in stage one, to the diagnostic second stage, to the eliminative third stage. De-projection serves, in other words, as a variety of referential coordinate transformation from the diagnostic first stage to the eliminative third stage, without risking the introduction of error that can come about, for example, by importing content from a different, dimensionally incompatible, non-complementary system of reference (see {10.5}).

When the fourth stage of de-projection is undertaken, however, the need to *interpret* a theory's original "intended meaning" introduces uncertainty. Since it is often possible to develop a plurality of potential interpretations of a theory's intended meaning or purpose, we have no assurance that a given interpretation is the best, is the most appropriate, or is unique. Nonetheless, if a proposed interpretation complies with the metalogical constraints upon reference identified in stage one, we are assured that such an interpretation is free of projection, and in this important sense, we can be confident that the proposed interpretation cannot be dismissed as incompatible with the theory's presuppositional structure.

Although the formal, tautological character of the first three stages of de-projective analysis cannot lead to new information not already contained in a given theory under analysis, and yet despite the fact that nothing new is introduced through its application, de-projective analysis allows us to become aware of what a given projective theory *cannot mean*: If such a theory has a salvageable meaning, its meaning cannot, in principle, be what its projective content putatively and deceptively attempts to express, which reflective analysis shows is necessarily devoid of meaning.

---

<sup>214</sup> An earlier detailed discussion of the tautological nature of de-projection may be found in Bartlett (1970, Section 1.4).

In this specific sense, de-projective analysis accords with the fundamental goals of a “*negative science*”: It provides a method that negatively defines the domain of possible sense of a theory by recognizing its metalogical horizons, which, if one attempts to transgress them, leads to meaninglessness and delusion. Although strictly speaking no new information is acquired, the de-projective analysis of a theory leads to an explicit recognition of a theory’s horizon of possibility and meaning, its metalogical horizon.

As we noted earlier in {12}, the elimination of meaninglessness is one of the central objectives both of formalized logics of meaning and of formalized logics of meaninglessness. As we saw in that chapter, some criterion or set of criteria of meaning is required in order to restrict a formal logic to meaningful statements and to exclude those that are not. But none of the proposed formal logics of either kind has developed, or justified, such a criterion or set of criteria. The method of de-projection accomplishes this end through the application of referential consistency as a criterion of meaning, which provides a systematic means to eliminate and exclude projectively meaningless theories, positions, and concepts, and, as a consequence, to eliminate and exclude their secondary expression in rational discourse.

### 15.3 Applying the method of de-projection

As noted earlier in this chapter, as a heuristic method, de-projection, at least at present, must rely upon a human metalogical analyst’s reflective skills and the ability to recognize and identify a theory’s fundamental metalogical presuppositional structure. The conceptual discipline, care, intellectual openness, initiative, and methodological rigor that are indispensable in this task can be especially challenging to cultivate and to maintain when habitually accepted ways of thinking lead the analyst to confront the counter-intuitiveness of his or her own attempts to stay the course throughout a difficult de-projective analysis. One of the most perceptive and articulate descriptions of this potentially demanding work was given by mathematician George Pólya in connection with the challenges sometimes posed by the method of indirect proof, or *reductio ad absurdum*. De-projection is not *reductio ad absurdum*, and yet the intellectual demands placed on the de-projective analyst can be similar. Here is Pólya’s description: Let us suppose that

We wish to prove that it is impossible to fulfill a certain condition, that is, that the situation in which all parts of the condition are simultaneously satisfied can never arise. But, if we have proved nothing yet, we have to face the possibility that

the situation could arise. Only by facing squarely the hypothetical situation and examining it closely can we hope to perceive some definitely wrong point in it. And we must lay our hand upon some definitely wrong point if we wish to show conclusively that the situation is impossible....

It must be confessed that “*reductio ad absurdum*” as a means of exposition is not an unmixed blessing. Such a “*reductio*,” especially if it is long, may become very painful indeed for the reader or listener. All the derivations which we examine in succession are correct but all the situations which we have to face are impossible. Even the verbal expression may become tedious if it insists, as it should, on emphasizing that everything is based on an initial assumption; the words ‘hypothetically’, ‘supposedly’, ‘allegedly’ must recur incessantly, or some other device must be applied continually. We wish to reject and forget the situation as impossible but we have to retain and examine it as the basis for the next step, and this inner discord may become unbearable in the long run.

Yet it would be foolish to repudiate “*reductio ad absurdum*” as a tool of discovery. It may present itself naturally and bring a decision when all other means seem to be exhausted.... (Pólya, 1971/1945, pp. 165, 168-169)

Pólya claims that applying the method of *reductio* “may become very painful indeed for the reader or listener”; my point here is that such pain must also sometimes be tolerated by the de-projective analyst. Where Pólya’s description speaks of the need to make recurrent use of such words as ‘hypothetically’, ‘supposedly’, and ‘allegedly’, we have already found that in de-projective analysis it is necessary to make repeated use of the phrases ‘putative meaning’, ‘purported intention’, etc. As in an indirect proof, in de-projective analysis we are caught in the situation in which we must deal, on the one hand, with a theory, position, or concept that is frequently used *as though* it were self-evidently and undeniably meaningful—so much so that to question it is may be experienced as “ridiculous,” “strained,” or “unnatural.” On the other hand, as de-projective analysis proceeds, we wish to apply strict standards of meaningfulness and must sometimes make a self-conscious effort to disregard the seductive nudges of conventional habits of thought and habitual language use. In both the method of *reductio ad absurdum* and the method of de-projection, we must deal with impossibilities, and yet in de-projection

we must attempt in some way to conceptualize these in order to salvage what meaning they may potentially have: This effort can not only be a mental strain, but, because we are dealing with metalogical impossibility in the form of projections, this effort can also be a cognitive impossibility. The experience, which Pólya calls “inner discord,” may at times seem extremely awkward, with one foot planted on firm ground, and the other on an unstable, leaking boat—while trying to maintain one’s balance.

In Part III of this book, we shall be in a position to apply the metalogic of reference to a range of theories, positions, and concepts whose metalogical analysis can admittedly be sometimes challenging, and certainly counter-habitual. If anything, the passage I’ve quoted above from Pólya communicates that intellectual endurance of a certain kind is a prerequisite when applying the method of *reductio ad absurdum*. A similar mental disposition is also, we shall find, indispensable when applying the method of de-projection. To bear up under the challenges of metalogical analysis calls, in the author’s experience, as much for *patience*, *stamina*, and *persistence* as it does for a special set of intellectual skills.<sup>215</sup>

#### 15.4 De-projection and framework self-enclosure

In {4.10}, the concept of self-enclosure was introduced to describe the way in which a system of reference comprises a closed system despite its absence of determinate boundaries. We noted that philosophical systems frequently have the property of self-enclosure in the related sense in which they form “sets that are closed upon themselves.” We observed that self-enclosure may characterize various kinds of systems, among them the frame of reference established in this study by its reflexive, vertical, non-ordinal frame of reference developed on the level of transcendental, maximum theoretical generality. We saw that from the standpoint of such a referential system, reference “outside” that framework is impossible since any reference will, by definition, presume that reference frame. We also noted, in a parallel fashion, that, for phenomenologists, experience possesses no boundaries; whatever is experienced becomes part of the self-enclosed field of experience.

In {10}, we continued this analysis, reaffirming, on a metalogical level, that the reference frame which comprises the subject of our study is itself self-enclosed: It is not theoretically possible to “go outside it” because, in addition to the reason given in the previous paragraph, it comprises a reference frame for the study of all frameworks of reference, including, reflexively, its own.

---

<sup>215</sup> This set of skills is described in detail in Appendix II.

In the last chapter we observed that a reference frame establishes a field in terms of which possible objects of reference are identifiable, and again saw how such a field is self-enclosed: Although this field possesses Leibniz boundaries that potentially may be extended indefinitely, yet such a referential field has no delimited boundaries, no boundary lines of demarcation, and is in this sense unbounded. However, as we saw, such a field does have “reactive” boundaries which are made evident by the fact that, when transgressed, they result in projections, undermining possible reference and meaning. Those reactive boundaries establish the metalogical horizon of a frame of reference.

The present chapter has developed de-projection as a method that identifies and affirms the metalogical boundaries of a given system of reference: The method is designed both to comply with and to enforce that system’s metalogical horizon, and in the process show that references which such a system makes possible must accord with the parameters that constrain possible references from its standpoint. In so doing, the method of de-projection establishes an explicit relation of *congruence* between a theory’s scope of possible reference and the metalogical parameters of constraint of its reference frame. De-projection discloses and enforces this fundamental congruence, and in this respect the method expresses one of the fundamental meanings of the term ‘tautology’.<sup>216</sup> In recognizing the metalogical boundaries of a theory’s frame of reference, and in enforcing that theory’s compliance with those boundaries, de-projection results in a self-referential acknowledgment of tautological equivalence.

This is the relation of tautological equivalence that obtains between (a) the preconditions of possibility described in the first stage of de-projection, and, as a consequence of stages two and three, (b) the resulting “de-projected” theory’s referential claims. This is a form of tautological relation that we shall call ‘equipollence’—from the Latin *aequipollent*, meaning “of equal value.” The term ‘equipollence’ has been used in the history of logic in a variety of contexts, dating back to the medieval logicians, and continues occasionally to be found, e.g., in deontic logic.<sup>217</sup> In its frequently used meaning in formalized

---

<sup>216</sup> Husserl similarly recognized this property of the structure of tautologies that I’ve termed ‘congruence’, fittingly calling it ‘autoconcordance’. See Husserl (1929, Appendix III) and Bartlett (1970, Section 1.6, pp. 121-128).

In the French translation, this is clearly expressed: “tautologies ... sont en quelque sorte des systèmes de la conséquence *qui se suffisent à eux-mêmes*, qui n’ont besoin d’aucune prémisses en dehors de leurs prémisses posées” [tautologies are in a sense systems [resulting] from the consequence that *they suffice in themselves*; they do not stand in need of any premise apart from their posed premises] (Husserl, (1965/1929, Appendice III §4 p. 432, italics added).

<sup>217</sup> For a detailed discussion, see Prior (1955a, esp. pp. 134ff; also pp. 14, 77, 147, 186-188, 220).

logics, the term has come to mean the interderivability of two expressions.<sup>218</sup>

To make this clear and concise, consider the relation between, on the one hand, the metalogical preconditions of reference,  $M$ , of a given theory,  $T$ , and, on the other hand, the theory  $T$  itself.  $T$  metalogically entails  $M$ , while  $T$ 's structural/systemic presuppositional structure is tautologically described by  $M$  (to hold  $T$  is, tautologically, to hold  $M$ —again,  $M$  supplies no new information not already ingredient in  $T$ ). We've called the relation between an accurate metalogical description of a system of reference (here the system of reference provided by  $T$ ), and the system itself, 'a relation of equipollence'. To give this relation a special name is to call attention to the intimate tie between a description of the referential preconditions of a frame of reference, and the reference frame so described. Although the foregoing is a metatheoretical observation, it bears some similarity to mathematician Ernst Zermelo's remark: "A definition may very well rely upon notions that are equivalent to the one being defined; indeed in every definition *definiens* and *definiendum* are equivalent notions...." (Zermelo, 1967/1908, p. 190).

The *equivalence of structure* of a system's parameters of referential constraint and the scope of the referential capacity of the reference frame under analysis is a consequence of the equipollence relation between the two.<sup>219</sup> This equivalence of structure as examined from a metalogical point of view determines what a theory can, in principle, refer to and mean.<sup>220</sup>

In disclosing this equivalence of structure, the formal structure of the method of de-projection is reflexive, possessing the property which in previous chapters {3, 4, 9, 14} we've called 'recurvature': Its structure is self-enclosed in the sense in which transcendental argumentation exhibits a special form of circularity ({8.4, 8.5, 9.1}), and in the topological sense of possessing no "outside."

From this theoretically abstract and general vantage point, we recognize that any frame of reference has its own metalogical horizon. De-projection supplies us with a means of recognizing that horizon, and the method compels our rational asset to comply with and to enforce the boundaries which that horizon establishes. De-projection leads to the result that the set of references that are *possible* from the standpoint of a given frame of reference is

---

<sup>218</sup> See, for example, Reichenbach (1947, pp. 107ff).

<sup>219</sup> Mathematician and logician Bernard Bolzano (2004/1851) employed a parallel use of the term 'equipollence' in connection with infinite sets when they meet the condition (as translated into English by Steve Russ) of "having exactly the same kind of construction" (but literally: "as having exactly the same origin") ("*die ganz gleiche Entstehungsart haben*").

<sup>220</sup> For further discussion of relations between tautological equivalence and equipollence, see Bartlett (1970, Section 1.4).

*referentially closed*:<sup>221</sup> No attempt, from the standpoint of a given frame of reference, to refer beyond that frame's referential capacity can succeed; any such attempt instead results in incoherence, in meaninglessness, and in beliefs which we shall find are, upon reflection, delusional and fundamentally deceptive.

### 15.5 *Reprise économique*

Let us insist upon a point which is perhaps already clear to everyone; but it is very important and, therefore, it should be completely clear.

– Pólya (1971/1945, p. 181)

Because the central principle of de-projection *is* very important to communicate clearly and effectively, let us consider an alternative approach to explanation before bringing this chapter to an end. It may be the case that for some readers talk of the “tautological nature of de-projection” and of “equipollence” does not fit in well with the furnishings of their minds, and may not provide them the satisfaction of having comfortably understood what the author has been trying to communicate.

When translated into our present framework, an informative concept drawn from economics may be put to clarifying use in this chapter's description of the method of de-projection. An economic system is said to be “*autarkic*” if it is *self-sufficient*—if the system can function without exchanges through trade or assistance from outside the system. ‘Autarky’ comes from the Greek αὐτάρκεια (*autarkeia*), meaning “self-sufficiency.” *Autarkeia* in turn derives from αὐτο-, meaning “self,” conjoined with ἀρκέω (*arkein*), “to suffice.” (‘Autarky’ should not be confused with ‘autarchy’ or ‘autocracy’, which have an entirely different meaning, i.e., government by a single and absolute ruler.)

Rarely has the concept of autarky been employed in discussions of logical systems, although it can be directly applied to many of them. One of the rare uses of the concept in a logical context was made by Hungarian mathematician George Pólya in his discussion of the syllogism. The following passage describes the logical structure of the syllogism in terms of its autarky—with

---

<sup>221</sup> Compare in formalized semantics the concept of a semantically closed language: “a language within which, as we are supposed to be able to do in ordinary English, we are capable of saying anything linguistically expressible...” (Berto, 2007, p. 232). Such a language is semantically closed in the tautological sense in which expressibility-in-that-language determines the scope of what it can be used to say.



much the same meaning as the author has been seeking to convey in connection with the logical structure of the method of de-projection:

The inference of a demonstrative syllogism requires nothing from outside, is independent of anything not mentioned explicitly in the premises. In this sense, the syllogism is *self-sufficient*: nothing is needed beyond the premises to validate the conclusion and nothing can invalidate it if the premises remain solid.<sup>222</sup>

This “self-sufficiency” or “autarky” of the syllogism is, perhaps, its most noteworthy feature. Let us quote Aristotle himself: “A syllogism is a discourse in which, certain things being stated, something other than what is stated follows of necessity from their being so. I mean by the last phrase that they produce the consequence, and by this, that no further term is required from without in order to make the consequence necessary.” (Pólya, 1968/1954, vol. II, pp. 112-113)

The method of de-projection possesses this property of autarky: The relation between the description of what is already ingredient in a system of reference—ingredient as a precondition of its referential capacity—and the system itself, is tautological, as we’ve observed, again in the original Greek sense of “repeating what has been said” (which I am again deliberately doing here).

To revert to our previous terminology which may now perhaps be clearer: The formal structure of the method of de-projection is *tautologically repetitive* in the sense that an accurate description of the preconditions of reference of a given theory bears a tautologous relation to that theory’s set of structural/systemic presuppositions that must be satisfied in order for reference to objects in that framework to be possible. As analysis moves from stage one to stage two to stage three of de-projection, nothing external to the theory being analyzed is introduced; the process of de-projective analysis is *self-sufficient*. No extraneous, no imported, information is introduced, but there is only an affirmation, compliance with, and enforcement of what is already metalogically entailed by the theory itself. A relation that I’ve called ‘equipollence’ is established between the set of preconditions of reference examined in these

---

<sup>222</sup> This same “self-sufficiency” was, perhaps unknown to Pólya, noted by Husserl decades earlier. See note 216.

three stages of de-projection, and the theory's reference frame. Nothing new is introduced that is not already included in that framework. The method of de-projection is, in a word, *autarkic*.

# 16

## Self-validation

### 16.1 Philosophical routes to certainty

Human error, the human propensity to make mistakes, to believe in delusions, to live as organisms prone to a multitude of forms of fallibility—these inexorable and unrelenting facts about our present cognitive and psychological constitution have motivated reflective minds to seek for ways to overcome these very human shortcomings.

For philosophers, scientists, and mathematicians, there has been no more important shared objective and value than the pursuit of truth. Over the millennia, many methods for reaching truth have been proposed and developed. In their very distinctive and diverse ways they aim ideally to supply routes to *certainty*, to results about which we can be *completely confident*—results that are *beyond fallibility*, beyond mere beliefs, speculation, and an animal's habitually formed opinions and expectations. Much of the history of the three disciplines can be understood from this perspective.

In the effort to develop forms of proof, demonstration, and justification that provide *completely reliable certainty*—a degree of certainty immune to the passage of time and impervious to the acquisition of new information—philosophy, natural science, and mathematics have centered attention on two approaches. Each of these applies its own special kind of standard of judgment. One approach is characteristic of mathematical proof; it is deductive. It is common to contrast the deductive approach of mathematics with the inductive method of science, but the inductive, hypothetico-deductive approach of science does not presume to lead to completely reliable, incontrovertible *certainty*, which is our topic here. The hypothetico-deductive approach is always, in principle, open to revision in the light of new information integrated in new theoretical models.

Instead, the second approach to unqualified certainty, which one frequently finds employed both in philosophy as well as in natural science, we'll call 'evidential'. Since much of the interest of the metalogic of reference is to establish a strongly compelling meta-framework that supplies the means to reach results that are certain, let us look at these two avenues to certainty, the

deductive and the evidential, more carefully.

The deductive approach is already familiar: In its strictest form, found in formalized systems, there are two main methods, the axiomatic and natural deduction. Both approaches have obvious weaknesses: The axiomatic approach begins by stipulating a set of postulates that must be hypothesized; the conclusions validly derived from that set of postulates, or axioms, are certain, but there remains the question whether the axioms themselves are also certain. Beyond stipulating or hypothesizing those axioms, there is little that can be called upon to *demonstrate* that their truth is certain. No matter how intuitively obvious a given axiom may seem, as in the famous case of Euclid's parallel postulate, a time may come, as in fact it did for Euclidean geometry, when good reasons become apparent that the axiom should be questioned.

Similarly, the approach of natural deduction formulates rules of derivation that are designed, to the extent that this is possible in a formalized manner, to mimic or embody forms of reasoning that accord well with what feels natural and intuitive to human logicians. The results derived through the valid application of such rules of natural deduction can legitimately be judged to be certain, but, again, it is open to question whether the *choice itself* of the rules can be established with certainty.

Beyond these formalized approaches that seek to yield results that are certain, there is moreover the role of *evidence* in proofs, justifications, and demonstrations. There are numerous meanings that the terms 'evident' and 'evidence' have acquired over the centuries. What concerns us here are those strictly focused meanings in which *evidence* is linked to *proof*.

The root-meaning of both terms derives from the Latin *evidentem*, meaning "clear, obvious, perceptible," and the Latin *evidentia*, the quality of being "obvious" and "evident to the mind." When the two words 'evident' and 'evidence' came into British use in the 11th century, they were often intended to convey "constituting *proof*." By the late 13th century, the term 'evidence' began frequently to be employed in law, meaning information that can be relied upon *to establish facts*.

Evidence in these limited and special senses supplies the basis upon which, in a variety of ways and contexts,

- ◆ a fact is established
- ◆ a justification is provided to support a conclusion
- ◆ information is gained that determines whether a proposition is true
- ◆ a source of evidence is found for judgments that are considered, in light of that evidence, to be "obvious" and "convincing," so

much so that such evidence leads to the meta-level decision to consider such judgments to be “indubitable,” “conclusive,” and therefore “certain”

Evidence that can achieve unshakeable results is of course *ideal evidence*—evidence that is raised on a high dais to such a degree that it may only infrequently be encountered in real life. Nonetheless, evidence in this sense defines an ideal, one which we may acknowledge is not often realized.

In this context pertaining to “ideal evidence,” philosophers have proposed a wide variety of views concerning such notions as

- ◆ indubitability
- ◆ self-evidence
- ◆ infallibility
- ◆ incorrigibility
- ◆ immunity to challenge

Generally, a judgment is considered to be “indubitable” if it is not possible to doubt its truth; it is considered “self-evident” if the evidence is so strong that it can admit of no other response than its acceptance. A judgment is regarded to be “infallible” if it is not possible, in one or more implicit or explicit senses of ‘possible’, for it to be false. A judgment is “incorrigible” if it is not possible for anyone to show that it is false. A judgment is regarded as “immune to challenge” if, however it might be challenged, such a challenge will inevitably fail. —Clearly, all of these notions, which wish to bring reliable certainty within grasp, would require substantial elaboration before any one of them is capable of offering a well-formulated standard of certainty.

Philosophical discussions and controversies that have had this aim have, as one might realistically expect, led to no conclusive, certain results. Philosophical arguments relating to evidential certainty have themselves failed to reach conclusions possessing the high degree of certainty that can be called ‘completely reliable certainty’.

Nevertheless, let us try, only briefly, to learn something from a few of the most relevant past efforts.

## 16.2 Distant connections with self-validation

A number of philosophical approaches have been proposed that either describe ways that certainty can potentially be achieved from a basis in evidence, or attempt to develop a method capable of reaching such certainty. Those that are of particular relevance to us in the present study include the

following five. They have been selected here for discussion—but without digressing to evaluate them critically—for two reasons: because all claim to be based on, or by themselves are thought to comprise, forms of evidence that lead to certainty, and because all express or imply an intrinsic acknowledgment of framework relativity.

- (1) Privileged introspection, often also called ‘privileged access’
- (2) Operative logic
- (3) Self-verification
- (4) Self-sustainment
- (5) Self-validation

*Privileged introspection* is the view that there are some things about which a human subject *cannot* be wrong. If there are things of this kind, then they are, literally, “incurable”—that is, they are not subject to possible correction (corrigibility). A possible example of this: I cannot be said to be wrong when I sense my own pain. Philosophical approaches that locate the certainty of judgments in introspective, immediate, so-called ‘immanent’ experience (e.g., Husserlian phenomenology) often claim this high degree of certainty. There are obvious potential objections to the claim that through introspection incontrovertible truth can be reached, given that the very notion of “incontrovertible truth” *means* that “no other can contest it”; but if others do not possess the same or compatible introspective access to that alleged truth, it cannot be established as true by an open community of thinkers.

*Operative logic* was developed by German logician Paul Lorenzen (1969, 1969a). Operative logic provides a “dialogue-operative” logical proof technique in which proof strategies reflect the *de facto* shared rhetorical basis of disputants relative to a universal audience.<sup>223</sup> The “shared rhetorical basis” of disputants refers to their shared standards of proof and the principles that govern their openness to persuasion. In Lorenzen’s operative logic, proof strategies are, so-to-speak, “operatively determined” as a function of these shared standards and principles. A logical framework is then operatively established in terms of the context provided by these shared tenets of reasoning. Relative to that operative framework, a valid logical proof obtains its level of certainty in the reflexive sense in which “the method is identified with its own result” (Lorenzen, 1969, p. 89). The certainty obtained receives its justification because, from the self-limiting standpoint of the operative framework, *there is, in fact, no appeal* to an alternative basis.

---

<sup>223</sup> On the concept of universal audience, see Perelman & Tyteca (1971).

*Self-verification* was proposed by Hintikka (1965/1962) in his interpretation of Descartes' *cogito* argument. Hintikka identified a category of what he called 'existentially inconsistent sentences': For example, the sentence "de Gaulle does not exist" is existentially inconsistent if de Gaulle himself were to utter it. Asserted by de Gaulle, the sentence becomes performatively self-referentially inconsistent (discussed earlier in {6.2}). "I do not exist," but asserted by *anyone*, is existentially inconsistent. In Hintikka's view, the *negation* of existentially inconsistent sentences like the one just mentioned "verify themselves" when they are asserted. Descartes' *Cogito, ergo sum*, Hintikka argued, is in this sense *existentially self-verifying*. As Hintikka expressed this: Descartes' *cogito* "... refers to the 'performance': (to the act of thinking) through which the sentence 'I exist' may be said to *verify itself*" (Hintikka, 1965/1962, p. 62, italics added). Self-verification, understood in this way, suggests one route to the degree of reliable certainty that concerns us. Unfortunately, Hintikka did not develop a generalized concept of self-verification beyond its existential application to Descartes' *cogito*.

In another work, Hintikka did, however, propose the related notion of *self-sustaining* claims, which were described in {11.5}. Such claims express truths which, he argued, possess "*immunity to certain kinds of criticism*": They are "truths which we *can* know without making use of any factual information" (Hintikka, 1962, p. 37), that is, truths we can recognize purely in an internal manner, in terms of what we might call their 'logical context'. This may be the logical context entailed by a particular argument or, by extension, the logical context established by a theory or position. Such truths as can be known in this context-relative way are, as Hintikka expressed this, "immune" to criticism because they are "self-sustaining"—as long as their logical context is presumed by someone claiming them, their validity cannot be challenged by a critic.

I shall ... adopt or, rather, coin, a few new terms.... [I]nstead of valid sentences I shall speak of *self-sustaining* sentences. Whenever an implication " $p \supset q$ " is self-sustaining, I shall say that  $p$  *virtually implies*  $q$ . If  $p$  virtually implies  $q$  and vice versa, I shall say that  $p$  and  $q$  are *virtually equivalent*. (pp. 31-32).

Hintikka employed the notion of self-sustaining sentences as a way of showing, by means of accepted rules or other techniques, that "a sentence can be made true by the sole means of internal criticism" (p. 36).

To translate Hintikka's suggestion into terms now familiar to us: We might say that he appears to describe in general terms a form of logical

analysis that *situates itself internally*—from within a given framework—and then *applies criticism internally* by making recourse only to the information, logical principles, and inferences that can be made by employing those principles, all of which are implicitly given in that framework. In this sense, a *self-sustaining claim* is one which logically follows within a given frame of reference, which we can know without recourse to additional information not already given, and which is therefore immune to criticism from anyone who both employs that framework and who would deny that claim.

It is appropriate to call this kind of internally situated analysis and criticism ‘*pragmatical*’: Like pragmatical self-referential argumentation (see {6.2}), Hintikka’s self-sustaining claims can be shown to be *logically inescapable* (although he did not use this expression) by showing that when an individual employs a certain framework while denying its logical consequences, he or she becomes pragmatically (or performatively) self-referentially inconsistent. The conclusion of such an argument is *certain* in the strict sense that interests us here. But such a conclusion is certain only given the following important and honest qualification, perhaps inadequately emphasized by Hintikka: “The applicability of our results may ... be said to presuppose a certain amount of rationality in the people whose attitudes are being discussed” (p. 38). We shall have occasion later on to discuss the central issue of rationality in detail.

Before continuing to the fifth and last potential route to certainty mentioned earlier, self-validation, we should recognize that the first four approaches—privileged introspection, operative logic, self-verification, and self-sustainment—are all of them *agent-centered*. That is, they are all pragmatically or performatively based; they all involve operations performed by someone, whether he or she introspects; engages in “operative dialogue” with fellow disputants; actively makes claims which, through the very performative act of making those claims, self-verifies the truth of those claims; or engages in internal performative criticism that culminates in self-sustaining claims.

In contrast to these agent-centered approaches that provide potential routes to reliable certainty, self-validation, as we shall see, functions on a general, purely conceptual level.

### 16.3 The “Worm of Ouroboros” logic of self-validation

We can see that what we need is ... a ground such that its “logic,” rather than being presupposed or stumbled on in an imaginative way, can be stated autonomously within a do-



main in which validation of categories includes that of the ground and so is self-validating.

– Klaus Hartmann (1966, p. 236)

It is one thing to claim that one can derive completely certain results from an individual's performative acts of reasoning (or its linguistic expression), whether these results are obtained through privileged introspection, operative logic, self-verification, or self-sustaining claims. But it is a very different thing to claim that one can derive results that are certain on the level of *possible meaning* and *possible knowledge*. The type of evidence appealed to by the former is of an altogether different kind and level than the evidence required by the latter.

When these two approaches are discussed, it has been customary and appropriate to contrast concrete applications of reasoning, on the one hand, with an analysis of their transcendental basis, on the other. As developed in previous chapters, that “transcendental basis,” as studied from the standpoint of the metalogic of reference, consists of preconditions that must be satisfied in order for referential consistency, and therefore meaning, rational coherence, and intelligibility, to be *possible*. To disclose and demonstrate such metalogical preconditions is a task that requires its own variety of logic.

Self-validation, as developed in {4} and {11}, incorporates a logic that is at once simple and yet potentially elusive. In this section, it will be useful to re-state and elaborate the principles that govern that logic.

As we've seen in the two preceding chapters, de-projective analysis renders explicit the implicit constitutive structure of a reference frame, and, relative to the structural/systemic presuppositions of that frame of reference, de-projective analysis makes strongly compelling the acceptance of and compliance with the metalogical horizon that expresses that framework's parameters of referential constraint. From the standpoint of the reflexive, vertical theory we are developing, we've seen that claims that seek to transgress a framework's horizon are projective and undermine their possible meaning.

A proof that a certain claim is projective demonstrates that the claim stands in conflict with the referential preconditions of its own possibility. We've observed that such a proof is “self-validating”: it cannot *not* be accepted without undermining itself on the metalogical level, on the level of its own possible meaning. To reject the conclusion of such a proof leads to the special variety of self-referential inconsistency we have now studied in detail, the projective form of self-referential inconsistency that undermines its own possible meaning.

The logic involved in de-projective analysis bears a surface similarity to Hintikka's remarks concerning self-sustaining sentences: In the de-projective analysis of a given theory, we cannot isolate a group of "premises" that do not already have built into them the "conclusion" we seek to establish: that is to say, the preconditions without which the theory's referential capacity is impossible are tautologically *built in*; a description of those preconditions is tautological in the sense that it adds no information already implicit in the theory under analysis. Where Hintikka's self-sustaining claims are logically entailed by a proposition or a position, metalogical preconditions are entailed by the very *possibility* of a theory's, a position's, or a claim's referential capacity, and hence its possible meaning. From Hintikka's point of view, for an individual to utter or affirm a given proposition is for him or her to be committed to whatever that proposition logically entails; here, instead, we find that for the referential capacity of a theory to be possible, compliance with its structural/systemic presuppositions is strongly compelled; not to comply with those presuppositions is to degenerate into projective meaninglessness. The theory itself metalogically entails the structural/systemic preconditions that de-projective analysis discloses. A recognition that this is the case is self-validating and tautologous.

The logic of self-validation leads to the pair of results summarized in {11.4}: (i) a statement (hereafter, or sentence or proposition) that cannot be *denied* (i.e., negated) without projection is self-validating, while (ii) a statement that is projective is such that its *rejection* (not its simple negation) is self-validating. Through the rejection of a projection, we obtain a statement that is self-validating, one that cannot be denied without metalogical self-referential inconsistency.

Several philosophers whose thought we have discussed in previous chapters have pointed in this general direction: We recall {8.4} Nelson's claim that the task of metaphysics is to disclose the preconditions of empirical judgments; in the process, they are shown to be valid, but, he believed, any attempt to prove them is unavoidably "circular." Palmer then criticized this peculiar cyclical logic that characterizes transcendental argumentation and called it 'p-circularity'.

We may also recall from {13.4} Royce's argument against realism—that it denies its own presuppositions and leads to meaninglessness because "the whole context of the realm of valid or possible experience ... *is not independent of its own fragments*," so that when you deny these fundamental presuppositions, "*you struggle in vain to articulate your meaning*" (Royce, 1959/1899, p. 248). This "realm of valid or possible experience" comprises for Royce what we've called a 'closed system'.

Coming a few decades later, Quine's comment, quoted in {10.1}, returns to mind: "Our argument is not flatly circular, but something like it. It has the form, figuratively speaking, of a closed curve in space" (Quine, 1963/1953, p. 30), and Lorenzen's remark, "the method is identified with its own result" (Lorenzen, (1969, p. 89).

These intimations of the recurved, reflexive logic of self-validation have been impressionistic and suggestive, but not conclusive. Transposed into the terms of the present study, as we also noted in {10.1}, the circular nature of this logic receives what we might figuratively call its "Worm of Ouroboros" character from the metatheoretical fact that the metalogic of reference must in a sense presuppose its task is finished before it is begun—in other words, to make a beginning is already to presuppose the *possibility* of the intended enterprise, a possibility that the metalogic of reference seeks itself to demonstrate by formulating its own structural/systemic presuppositions. This reflexivity, rather than a fault of logical circularity to be criticized, is what we must expect when dealing with the self-enclosed, recursive nature of dynamic, general systems of reference. Framework relativity, in other words, is a recognition of the "closed circuitry" nature of the logic of self-validation.

I have sought to make clear that the systems-based structure of transcendental argumentation cannot be captured by a syllogistic, argument-based approach, but rather requires a logic that is capable of expressing the formal structure of the method of de-projection as inherently *reflexive*, possessing the property which in previous chapters {3, 4, 9, 14} I've called 'recurvature': Its structure is self-enclosed in the sense in which transcendental argumentation exhibits this special, legitimate form of systems-circularity, which provides the basis for the self-validation of de-projective analysis.

At the beginning of this chapter, I quoted a passage from a paper published long ago by Klaus Hartmann. In his discussion of the thought of Hegel, Hartmann, like the philosophers I've mentioned in the preceding paragraphs, noticed the same phenomenon of "circularity" that appears in attempts to engage in transcendental reasoning, and he, too, thought to apply the term 'self-validating':

We anticipate thought as a result when we "use" it in the process of categoreal development. Just as we can consider thought justified by the introductory argument of the *Phenomenology of Spirit*, so we can regard it *as self-validating in view of its end which is its ground.*

In short, the Hegelian proposal is to avoid the problem of a first stance by invoking circularity, not now in terms of

granting and reconstructing only, but in terms of a theory of categories whose justification is borne out by the result of the categorical doctrine itself. In other words, the principle of transcendental explanation itself can be grasped. This, however, in its domain, is the domain where its explanatory logic holds. Only the domain is now universal and all-inclusive in the sense that there is no division between ground and grounded left as one between a fixed stance and derivative results. The whole domain is self-grounding, self-validating. From “outside” it looks like a *petitio principii*, but only from outside. From inside we can demonstrate its *systemic* virtues. (Hartmann, 1966, p. 238, italics added)

#### 16.4 The irrefutability of self-validation

There is a clear sense in which the valid inferences made within consistent axiomatic systems or within consistent systems of natural deduction are “irrefutable.” The strictest meaning of the term ‘irrefutability’ comes from this application to consistent logical systems. By the expression ‘*irrefutability within a system*’ what I mean is the invulnerability within a system to any attempted disproof of valid inferences in that system such that the tools available for any potential disproof are limited to the proof-resources provided by that system. Otherwise, if postulates or rules of inference that are not authorized by the system in question are applied to it, the result is not “refutation,” but rather “*disagreement*,” which can pave the way for the development of one or more alternative systems, but does not in any sense constitute a refutation of inferences within the initial system. The assertion of a non-Euclidean parallel postulate does not in any meaningful sense “refute” Euclidean geometry; it may point to an alternative geometry in which conclusions may be derived that disagree with those inferable in the Euclidean system. In other words, assuming Euclidean geometry to comprise an internally logically consistent system, propositions that are validly derived in terms of that system are “irrefutable”: they are certain relative to the Euclidean system. This framework-relative meaning of ‘irrefutable’ applies directly to the logic of de-projection.

In de-projective analyses that self-validate, and provided the results of such analyses are reached in a manner consistent with the principles of the metalogic of reference, those results, relative to the theory (position or concept) under analysis, are *irrefutable* in the framework-relative meaning of this word. Given the observations we have made in previous chapters concerning

framework relativity, we are compelled to recognize, on pain of projective incoherence, that irrefutability in some unrestrictedly broad sense—“all-encompassing and autonomous of all frameworks”—is not, in principle, within possible meaningful grasp. No matter the system employed to prove a certain result, that result, tautologically, presupposes the referential and proof resources of that system, or of a system(s) equivalent to it.

In maintaining the irrefutability of self-validation we do, however, need to set down these provisos:

First, the de-projective analysis of a given theory, position, or concept requires a well-formulated, non-ambiguous starting point. The given theory’s fundamental constitutive structure must be explicitly formulated. This task is, on the one hand, *descriptive*, but it is at the same time also *stipulative*. The results of a de-projective analysis are based exclusively on the stipulative description of a particular theory’s structural/systemic presuppositions. The “Ouroboros” logic of self-validation is irrefutable only relative to the stipulative description that provides the basis for de-projective analysis.

Second, since we accept that the method of de-projection is heuristic in nature, not a method which we at present can translate into a potentially automated algorithmic procedure, the self-validating results of a de-projective analysis require our careful critical reflective scrutiny. As in a mathematical proof, it is possible to be misled into a premature, and perhaps mistaken, positive evaluation of the validity of a proof. Given this second proviso, in order to confirm the irrefutability of a self-validating proof of a de-projective analysis, reflective *confirmation* is required; once confirmed, the proof stands.

When these two provisos are satisfied, we may conclude that a confirmed self-validation is not subject to refutation: It is reliably certain to the same degree and in the same sense as is any confirmed mathematically proved result.

...

Toward the beginning of this chapter, we reviewed several ways in which evidence is variously relied upon—for example, to establish facts, the justification for a conclusion, information that supports the truth of a proposition, or judgments that, in light of such evidence, are considered to be obvious or convincing, perhaps even indubitable, conclusive, and certain. In the majority of such instances, the evidence relied upon is *external* to the facts, propositions, or judgments which, by virtue of that external evidence, are pronounced to be the case, true, or dependable.

We then listed some of the general—by and large historically vague and

inconclusive—terms used by philosophers who have claimed of some of their arguments and positions that they are “indubitable,” “self-evident,” “infallible,” “incorrigible,” or “immune to challenge.” We moved beyond these generally imprecise and elusive attributions of certainty and instead focused attention on a group of more specific kinds of evidence to which philosophers have appealed: These have included privileged introspection, operative logic, self-verification, and self-sustainment. We recognized that the evidence which all of these approaches rely upon in order to establish a basis that provides alleged certainty is *intrinsic*, *internal*, and *framework-relative*, and yet, also in all cases, these approaches are *agent-centered* and *performative*, and therefore subject to the “almost inconceivable hardihood” of philosophers, observed by Passmore ({6.5}), to resist the conclusions reached by other philosophers.

Rather than employing a pragmatist form of argument, the certainty acquired from de-projective analysis is also intrinsically based, but reached as a result of a metalogical, internal, framework-relative analysis of a given theory. The certainty acquired is not performatively basis, but rather follows necessarily from the constitutive structure of the reference frame studied. This is an altogether different variety and level of evidence, originating from within the very theory, position, or claim that is the subject of analysis.

The most commonly encountered, familiar form of evidence is external to a claim requiring evidential support; the kind of evidence which concerns us here is, instead, internal and metalogical, an acknowledgment and recognition of a reference frame’s self-validating coincidence with its own constitutive structure. This is an acknowledgement and recognition of what we might, extending a concept suggested by Lewin, call a metalogical assertion of a system’s ‘genidentity’,<sup>224</sup> or, as Husserl was quoted in the last chapter, we might call this a metalogical assertion of the system’s ‘autoconcordance’. More compactly expressed, the equipollent relation of metalogical preconditions to the reference system that is made possible through them is self-supporting, self-validating, and, from the standpoint of that system, irrefutable.

---

<sup>224</sup> As paraphrased by Reichenbach (1965/1920, p. 53): “[T]he principle of genidentity ... indicates how physical concepts are to be connected in sequences in order to define ‘the same thing remaining identical with itself in time’.” Kurt Lewin (1920, 1920a) developed this concept.

# 17

## Rationality: Rules of Admissibility

Modern heuristic endeavors to understand the process of solving problems, especially the *mental operations typically useful* in this process. It has various sources of information none of which should be neglected. A serious study of heuristic should take into account both the logical and the psychological background....

– George Pólya (1971/1945, pp. 129-130)

### 17.1 The two sides of rationality

Rationality has two complementary and interconnected sides: One side relates to the possible *conceptual rationality* of a theory, position, or claim, the conceptual rationality of a method or procedure, of a model of a process, of the rules of a game, etc. The other side of rationality relates to specific, often specialized, rational *cognitive abilities, dispositions, and traits* of human or other agents. These two sides are commonly interlinked when we say in very general terms that rationality requires both the conformity of an individual's beliefs and claims to knowledge with his or her justification for holding such beliefs and claims, as well as the conformity of his or her resulting thought processes, judgments, and actual behavior with those justified beliefs and claims.

As readers are by now aware, the chief focus of this work is conceptual rather than pragmatic and psychological. Yet in any general theory of rationality, both components, conceptual and psychological, play important roles. In this chapter, we shall need to discuss both, but in doing this, our scope of interest will be strictly delimited.

Every discipline has admissibility standards that act as regulative controls which are applied in judgments that determine acceptable research and proof. These standards of admissibility are often not explicitly stated, but tend to be revealed in the practical decisions that researchers make when they engage in research, formulate and publish their findings, and criticize the work of

others. The specialized admissibility standards that apply to acceptable research and proofs in mathematics can be stated rigorously, much more so than they can in connection with other disciplines. In relation to many other disciplines, the admissibility standards that are applied to philosophical research and to the conclusions to which philosophers arrive are considerably vaguer, less explicit, less consistent, and less agreed-upon by the different approaches to philosophy.

Clearly formulated *rules of admissibility* have been established and are routinely followed by courts of law; they include, for example, rules that set standards for the admissibility of evidence and that define limits which serve to exclude the introduction of irrelevancies. The admissibility and exclusion controls that regulate court practice exemplify in the clearest form prototypes of these two kinds of disciplinary restraints.

In both mathematics and in law, such rules of admissibility function in actual practice both to set down *criteria of acceptance*, as well as *criteria of exclusion*, and so we could, if we wished, distinguish *rules of admissibility* from *rules of exclusion*. However, for convenience in what follows, we shall assume that rules of exclusion form a constituent part of rules of admissibility.

In {11}, we discussed a wide variety of kinds of meaning, as well as a range of proposed theories of meaning. Similar to the nature of meaning, numerous varieties of “reason” and of “reasoning” have been distinguished during philosophy’s and logic’s long history, while the concept of “rationality” has accordingly been defined in a multitude of ways. Just as there may be no convincing and adequate single encompassing definition of ‘meaning’—that is, a definition of a *sufficient condition* that applies to every form of meaning—so is this apparently true in connection with the terms ‘reason’, ‘reasoning’, and ‘rationality’.

Furthermore, just as none of the criteria of meaning that philosophers have proposed have been strongly compelling in the sense introduced in this book, so is this also true of criteria of rationality. A wide array of criteria or standards of rationality has been proposed by logicians, mathematicians, psychologists, philosophers, economists, and others. It would take us too far afield here to review the many definitions of rationality that have been recommended. It is safe to say that none has to date achieved the high degree of *strongly compelling* status that we have previously defined—that is, so strongly compelling that its rejection is self-undermining on the level of *possible meaning*. Traditionally proposed criteria of rationality are subject to controversy, disputation, and rejection, and as a consequence none has yet to be universally embraced.

The reflective reader may immediately wonder if this last claim is in fact



true: Traditionally advocated standards of rationality, such as the principles of non-contradiction and excluded middle, may come to the reader's mind as principles that must surely be universally accepted; but these, like Euclidean geometry, have lost their persuasiveness in the wake of the development of alternative logics that reject or loosen them. Developments, for example, in game theory, behavioral economics, sociology, and the continued production of philosophical theories of rationality, when taken together in a synoptic view, make it clear that "rationality," if not a concept lost on Clifton Fadiman's "scrapheap of popular misuse," is at least a concept still in search of a definitive and universally persuasive meaning.

In connection with the subject of meaning itself, this study has developed and made the case for a *necessary condition* of meaning in the form of referential consistency. It is a criterion of meaning, as we have seen, that is *strongly compelling* in precisely the above-mentioned sense that its rejection is self-undermining on the level of *possible* meaning. In the present chapter, we shall continue the same path, defining a *necessary condition of rationality*, one which, like referential consistency as a criterion of meaning, is also strongly compelling.

## 17.2 Intelligibility and coherence

In the course of previous chapters, I have periodically and in passing made use of the terms 'coherence' and 'incoherence', and 'intelligibility' and 'unintelligibility'. These terms are closely tied to the concept of rationality that will concern us here.

*Coherence*, like *meaning* and *rationality*, is a concept claiming numerous alternative definitions. In this study, the concept of coherence has a specific and limited meaning: For our purposes, it can be defined in two senses, one positive and one negative: {8.3} developed the concept of structural and systemic presuppositions. We recall that such presuppositions are conditions upon which an object, an organized collection, or an interrelated system necessarily depends in order to "cohere" or to function as an integrated ensemble. Coherence in this sense refers to the fundamental organization of a totality that is responsible for its capacity to function in an integrated way, that is, to cohere as a system.

The second meaning of the concept of coherence can be defined negatively in terms of the result that comes about when a theory, position, or claim undermines its own possible meaning. As we have formulated this, when such a theory (henceforth also: position or claim) is projective, it becomes "incoherent": It becomes not only meaningless, but its meaninglessness is due to its

loss of coherence; because it is projective, it undercuts its own structural/systemic basis; its capacity to function as a coherent, integrated totality is undermined. Such a theory is in this sense “incoherent.”

The concept of *intelligibility* is somewhat different. It is an explicitly framework-relative concept that presupposes an underlying frame of reference in terms of which a theory’s “sense can be recognized.” In more psychological terms, a certain minimal level of “intelligence” is presupposed for the meaning of a theory to be apprehended. That minimal level of intelligence supplies the reference frame in terms of which a theory’s meaning *can* be understood. The common implied meaning of the term ‘intelligible’ is just this: A concept or expression in language, for example, is “intelligible” if, given a minimal level of intelligence in a perceiving agent, that agent can “make sense” or “recognize the sense” of that concept or expression. Intelligibility is framework-relative in this way. It was in such a context that Polányi (1959, pp. 21-22) observed: “...the sender of [a] message will always have to rely for the comprehension of his message on the intelligence of the person addressed.”

The metalogic of reference claims that without referential consistency meaning is impossible, and, hence, intelligible experience, too, becomes impossible. If the parameters of constraint of a given reference frame are undermined, it can no longer function coherently, and so cannot supply the necessary basis for the recognition of meaning. That which is metalogically self-undermining—i.e., projective—obstructs the possibility of meaning, and hence undermines a concept’s, expression’s, or theory’s intelligibility. Intelligibility, then, presupposes two fundamental things: Both a framework in terms of which something may be recognized as meaningful, and the meaningfulness of that which is in that way recognized.

When we say that projections result in conceptual incoherence and unintelligibility, what is putatively intended by such projections is recognized through reflective analysis to be devoid of possible meaning. A framework’s metalogical horizon delimits the scope of its referential capacity, beyond which lies only meaninglessness, which is to say *unintelligibility*. As this was expressed in this book’s Introduction, the metalogic of reference seeks to develop an approach to philosophy that provides a constructive, definite, and conclusive basis that cannot *not* be accepted without fundamental and self-defeating *incoherence*.

### 17.3 Epistemological rationality

A necessary condition of rationality is now within our reach. Let us give this necessary condition a special name, ‘*epistemological rationality*’. Epistemo-

logical rationality as we shall define the concept stipulates two rules of rational acceptability; each rule is compound. One of these rules applies to the *conceptual* side of rationality, and the other to the *cognitive-psychological* side:

- (i) a frame of reference is presupposed that is non-projective, coherent, and provides a basis for the intelligibility of objects of reference that fall within its scope of reference, and
- (ii) a human or other agent is presupposed who is competent to employ (i), and for whom rationality and conviction are fused

The first of these compound rules forms a *conceptual* prerequisite for rational admission: A given frame of reference must be free of projection, systematically coherent, and must have the capacity to refer to a range of objects of reference. We shall call such a frame of reference '*rational*'.

The second rule requires, on the one hand, that a human or other agent possess adequate *cognitive* competence to use that frame of reference, and, on the other hand, that the agent's resulting rationality is *psychologically* tied to his, her, or its *conviction*, as we shall make clear in what follows. We shall call such an agent '*rational*'. The first of these rules should by now in this study require no further explanation. Let us therefore turn to discuss the second.

The second rule relating to the psychological component of rational admissibility requires that a human or other agent possess "adequate competence" to use a given frame of reference; this means that the agent knows how to use the coordinating reference frame in question. It may be a simple Cartesian coordinate system, or a more complex system of identifying reference that is presupposed by a theory of physics, or by a system permitting pattern recognition of a certain kind, etc.

Beyond this, when we say that the agent's rationality and conviction are "fused," this relates to our previous discussion of the "bifurcation of mind" in {1.1}. There, we called attention to "the rational bridge problem," the problem which, in the author's judgment, poses important challenges for humanity: The rational bridge problem refers to the fact, that, for a great many people, well-developed reasoning skills fail to carry over into their rational decision-making as expressed in their actual conduct. This absence of "carry-over" identifies the rational bridge problem.

In {1.2}, four prerequisites were described that must be satisfied in order for a philosopher, one who is firmly committed to his or her preferred beliefs, to change his or her mind. To recall them, they were:

- (i) an intellectual capacity conjoined by a psychological willingness both to consider a view not one's own, and to measure against that view one's own position through the use of agreed-upon standards of rationality and validity
- (ii) the supposition that there actually *exists* such a set of criteria that *is* agreed upon
- (iii) a mental constitution that enables the philosopher to cross the bridge of rationality, and, in making that passage, form a persisting, non-provisional, non-tentative *internal commitment* to the rational results that are reached: that is, he or she must be *convinced* by reason on a fundamental emotional and behavioral level in a manner that strongly influences future thought and conduct
- (iv) the mind of such a philosopher should ideally be the kind of mind that is both willing and wishes to discover universal, compelling truths. Without this overriding intellectual temperament and attitude, changes to a philosopher's mind tend to be confined to delimited, often fragmented, specialized, technical topics, limited in range of applicability, and falling short of the needs of a systematically integrated understanding.

(This "psychological profile" of philosophical rationality is intentionally abbreviated for our purposes here. For readers interested in a more complete description of the cognitive-psychological side of epistemological rationality, see Appendix II, "Epistemological Intelligence.")

For the greater part of the human race, rationality and conviction seldom go hand-in-hand. For most people, an uncritical and dogmatic self-assurance dominates their minds according to which their preferred beliefs are judged to be true in proportion to the degree of conviction they invest in them.<sup>225</sup> In

---

<sup>225</sup> See, e.g., Bartlett (2005, 2011) and Appendix II in the present study.

{6.5}, we saw how Passmore pointed his finger at the shakiness of what I've called 'the rational bridge', and recognized that philosophers are encumbered by an "almost inconceivable hardihood" to persist in their pet beliefs. With much the same implication, Johnstone similarly observed the "almost unlimited capacity of sophistry" of many philosophers to defend their positions to the last gasp. For much of humanity, a matter is true *because* it is believed.

The two rules of rational admissibility stipulated above have the effect of restricting both the range of conceptual tools that are *acceptable*, as well as the size of the group of human or other agents who are *capable* of using those tools.

And so when, in {11.4}, the concept of *metalogical rejection* was developed as a regulative principle to eliminate and disallow projections, the concept of metalogical rejection presumed *rationality in the sense defined in the present section*.

In characterizing transgressions of metalogical horizons as non-metaphorically *delusional*, we re-assert and comply with the above rules of rational admissibility. Those rules, to re-emphasize this, function in a regulative capacity both to restrict what is admissible, as well as to exclude what is not. Specifically, in de-projective analysis we exclude projective theories (positions and claims), and seek, when possible, to reformulate such theories in a respectfully sympathetic manner that eliminates and avoids projective delusion. These rules of admissibility block the *exemption* of beliefs that do not comply with criteria of referential consistency, and in doing this, they block and exclude beliefs that fail to satisfy the necessary conditions of rationality.

When these rules rational admissibility are satisfied, what remains are claims (theories or positions) which, on the conceptual side, are expressed in terms of non-projective, coherent frames of reference that provide a basis for the intelligibility of objects falling within their respective scopes of reference, and which, on the cognitive-psychological side, are affirmed by a human or other agent who is cognitively and psychologically competent in the rational sense that we've defined.

The perceptive reader of the previous chapters in this study will immediately see that if the two rules of admissibility that define epistemological rationality are violated, we would be left in a position that is entirely vulnerable to self-undermining claims that transgress metalogical horizons of possibility and meaning. We would have no coherent frame of reference to provide a basis for intelligibility, nor would we have the competence necessary for coherent, intelligible use of such a framework. It is in this sense that the rules of rational epistemological admissibility which we have described cannot *not* be accepted precisely because their rejection is self-undermining on the level of

possible meaning.

The metalogic of reference is chiefly concerned with the conceptual side of the rational relationship we have outlined. But the metalogic of reference possesses clear-cut applications to the practical formation, self-criticism, and maintenance of beliefs by human or other agents—but only when the foregoing rules of rational admissibility are satisfied. When I have used the phrase ‘strongly compelling’ in connection with the self-validating results reached by this study, these rules of admissibility are presumed to be satisfied.

In the philosophical literature, there is perhaps no more articulate passage that expresses the fundamental spirit and motivation of this chapter than the following, written by British mathematician William K. Clifford. Where Clifford urged the need to justify our beliefs by means of evidence, the focus in this chapter has been the strongly compelling nature of epistemological rationality.

No simplicity of mind, no obscurity of station, can escape the universal duty of questioning all that we believe.

It is true that this duty is a hard one, and the doubt which comes out of it is often a very bitter thing. It leaves us bare and powerless where we thought that we were safe and strong. To know all about anything is to know how to deal with it under all circumstances. We feel much happier and more secure when we think we know precisely what to do, no matter what happens, than when we have lost our way and do not know where to turn. And if we have supposed ourselves to know all about anything, and to be capable of doing what is fit in regard to it, we naturally do not like to find that we are really ignorant and powerless, that we have to begin again at the beginning, and try to learn what the thing is and how it is to be dealt with—if indeed anything can be learnt about it. It is the sense of power attached to a sense of knowledge that makes men desirous of believing, and afraid of doubting....

[I]t is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence.... It is wrong in all cases to believe on insufficient evidence; and where it is presumption to doubt and to investigate, there it is worse than presumption to believe. (Clifford, 1877, pp. 293, 295, 309)

*A horizon defines, from your present standpoint, not only how far you in fact you see, and how far you can possibly see, but it defines, given your frame of reference, the limits of what can possibly be meaningful. If you attempt in your claims to knowledge to go beyond that horizon while maintaining your present standpoint, you will trespass beyond the boundaries of possible meaningfulness.*





# PART III

## PHILOSOPHICAL APPLICATIONS OF THE METALOGIC OF REFERENCE

### Major Problems and Questions of Philosophy and the Philosophy of Science

Concepts which have proven useful in ordering things easily acquire such authority over us that we forget their earthly origin and accept them as unchangeable facts. They are then stamped as “necessities of thought”, “given *a priori*” etc. Such errors often make the path of scientific progress impassable for a long time. It is therefore by no means an idle pastime if we become practiced in analyzing the long-established concepts and in showing the conditions on which their justification and usefulness depend, and in detail how they have grown out of the givenness of experience. By this means their excessive authority is broken. They will be removed if they cannot properly be legitimated; corrected if their coordination [*Zuordnung*] with given things has been too careless; replaced by others if a new system can be developed that we prefer for whatever reasons.

– Albert Einstein (1916, p. 102, author’s translation)

This part of the book applies the heuristic principles developed in previous chapters to a group of problems and questions that has occupied philosophers for more than two millennia. We shall discuss, one at a time, a series of concepts that recur frequently in the thought and expression of philosophers, concepts that commonly attract and hold the unquestioned belief of perhaps the majority of philosophers. These concepts form the conceptual vocabulary for much past and present philosophical thought and writing. They include beliefs about reality, space, time, agency, causality, and the reflexive nature of the discipline of philosophy itself. Treating these items separately is in a sense

artificial, because no one's belief system is made up of separable items: It rather forms a unified and dynamic system that provides the foundation for the individual's thought and its expression. The need to examine these concepts separately is necessary, for we can only discuss one thing at a time, often at the price of removing it from its living, interrelated context.

The results we shall reach are frequently both counterintuitive and run counter to traditional philosophical thinking. In the analysis of each question and problem, a conclusion is reached which, it is claimed, cannot *not* be accepted without incurring the special variety of self-undermining referential inconsistency which the author has called 'projective'. Results validly reached in this way are strongly compelling, and demonstrate that much that has pre-occupied philosophers in the past is solvable in a permanent and determinate way.

To reach these ends, readers will be greatly assisted if they can achieve, to paraphrase Coleridge, a willing suspension of their habitual and favored beliefs, "awakening the mind's attention from the lethargy of custom."<sup>226</sup>

---

<sup>226</sup> Coleridge (1817, Chapter XIV).

## Methodological Recapitulation

In previous chapters, the development of the metalogic of reference has led to the formulation of a heuristic method with two objectives: On the one hand, its chief negative objective is to identify and to eliminate the special variety of self-referential inconsistency that we have called ‘projective’, a form of inconsistency that undermines the possibility of both reference and meaning. On the other hand, the method’s principal positive purpose is to reformulate projective theories, positions, and concepts in a manner that seeks, when this is possible, to conserve the genuine meaning they may have while their projective content is removed.

In Part III of this study we embark on a group of applications of the metalogic of reference to a selection of central problems and questions of philosophy. Before we begin, it will be useful concisely to recapitulate the procedure we shall follow in each application.

‘De-projection’ is the name we’ve given to the heuristic method designed to realize the above pair of objectives. The method proceeds by means of four stages: descriptive, diagnostic, eliminative, and corrective.

In order to apply the heuristic steps that comprise de-projection, we need to begin with a well-formulated, non-ambiguous starting point, one that formulates in an explicit manner the fundamental referential structure of the theory under analysis. On the one hand, this task is *descriptive*, but, on the other, it is inescapably *stipulative*. De-projective analysis and the strongly compelling results to which it leads are claimed to be valid only in relation to and in compliance with the strictly stipulated, defined parameters of referential constraint upon which a given analysis is based.

The *diagnostic* stage of de-projection informs us whether the purported objects of reference of a given theory can, in principle—given the parameters of constraint of that theory’s presupposed reference frame—comprise possible objects of reference, and then determines whether the theory’s intended forms of reference conflict with its metalogical presuppositions. In cases of such conflict, the intent is to make explicit that projections are involved, and to verify that such projections undermine their own possible meaning.

When the preceding diagnostic stage has established that the theory under analysis does in fact involve projective reference, we proceed to the *eliminative* stage of de-projection in which the strongly compelling rules of rational admissibility that we endorse lead to the rejection of a theory’s projective

content. We recall that to *reject* such projections is *not* to *assert their negations*: it is rather to disallow and dismiss the projective nature of the theory under analysis.

The final *corrective* phase of de-projection attempts to re-formulate such a projective theory in a manner that seeks to preserve its possible non-projective meaningfulness, in conformity with that theory's regulative constraints upon possible reference.

Each application of the method of de-projection is understood in what we might call a 'quasi-axiomatic' sense: Each application is to be evaluated as a "*closed proof*" —that is, the stipulative description of the parameters of referential constraint relating to the theory under analysis leads to a *self-contained demonstration*. If an alternative initial stipulative description were to be accepted, then of course other results would follow. The self-validating claim of each application is therefore entirely framework-relative in this closed sense. This is the same framework relativity that we see in any axiomatic system: If alternative axioms are stipulated, alternative results from them may be provable.

This approach to de-projective demonstration is employed precisely to circumvent the usual philosophical disputation and controversy that have beset attempts by philosophers to establish their conclusions in an unobjectionable manner.

As long as the basic descriptive starting point of a reflexive metalogical proof is left open to alternative formulations and alternative assertions, *any* conclusion whatever can be reached. This is no rhetorical sleight-of-hand—it is a simple recognition of the very nature of logical proof.

Finally, the confirmation that a de-projective analysis is self-validating requires reflective confirmation. Just as the application of the heuristic method of de-projection requires the reflective capability of some agent, so does the confirmation of the self-validating results reached.

In the chapters that follow, in order to avoid what could read as the plodding monotony of a series of recipe-oriented applications of the method of de-projection, I will not always call attention to the separable steps of de-projective analysis and will rely on the reader to distinguish these when desired as he or she proceeds.

As volcanologists say,

*“The deeper we go, the less we know.”*



# 18

## Ontology and the Metalogic of Reference

For although in a certain sense and for light-minded persons non-existent things can be more easily and irresponsibly represented in words than existing things, for the serious and conscientious historian it is just the reverse. Nothing is harder, yet nothing is more necessary, than to speak of certain things whose existence is neither demonstrable nor probable. The very fact that serious and conscientious men treat them as existing things brings them a step closer to existence and to the possibility of being born.

– Hermann Hesse (1979/1943, p. 2)

### 18.1 The need for philosophical perspective

**L**ike most inbred almost exclusively academic disciplines, philosophy has been sorely deficient in possessing a sense of humor about itself, about the degree to which some of its hallowed problems are trifling, trivial, and intellectually frivolous. It ought not to be excessively heretical to suggest that a small dose, even, of self-ridicule can occasionally be salubrious.

Perhaps for none among philosophy's questions and problems is this more true than the status of non-existent objects. Whether or not non-existent objects "exist," and if so, in what of many possible senses of the word 'exist' they do, has occupied a great many hours of philosophers' grave reflection and has filled not only many thousands of pages of solemn and earnest discussion, but, of course, also philosophically inevitable controversy.

If one accepts the long-heralded notion that the main general tasks of the discipline of philosophy are to make sense of the universe and to improve philosophical thought itself, concern over the ontological status of non-existent objects would seem to have very little significance. The problems and questions that relate to this concern bear the marks of artificial artifacts of ingrown theorizing about what, at base, is—in most attempts to urge a particular position concerning non-existent objects—an arbitrary decision of

logic, that is, a decision that cannot itself be justified logically but is merely an expression of a philosopher's or a logician's "taste" or "preference" reflected in his or her favored ontology. The preponderant ontological "taste" in recent decades has been to favor the view that only physical things "exist," and all other things, including fictions, abstract objects, etc., do not.<sup>227</sup>

Like so many of us, philosophers are prone to take themselves too seriously and to magnify the ultimate importance of problems that have become important almost entirely because they have become important to philosophers who have a taste for them. To take oneself too seriously is largely to lose a sense of perspective, and when perspective is lost, then there comes to full expression the very strong human propensity to inflate the value of what has merely come to feel important, and to cultivate an uncritical willingness to devote time, work, interest, and deadly serious analysis to subjects whose main substance lies in the attention that has been rallied to invest in them.

The question whether non-existent things exist is such a subject. It is not a subject simply to be mocked, for the question does possess some meaning that can be instructive—but, at the same time, there is a need for the perspective provided by a sense of humor—for what can elicit cosmic laughter more than the fact that a species would concern itself with whether, how, and in what ways non-existent things exist? In expressing the matter in this over-simplified way, I do not mean to suggest that the question ought cavalierly to be tossed aside as so much over-intellectualized fanfare over nothing, although *nothing* does clearly play a leading role; it is not a question that can be resolved bluntly and simply merely by a quick kick imparted to Dr. Johnson's stone—but it can, very nearly so.

Interestingly and also somewhat ironically, examining the ontological status of non-existent things will provide us with fundamentally important information about the general ontology of objects that—in a vast variety of different ways—*are*.

## 18.2 The ontology of non-existent things

Whatever may be an object of thought, or may occur in any true or false proposition, or can be counted as *one*, I call a *term*. This, then, is the widest word in the philosophical vocabulary. I shall use as synonymous with it the words unit, individual, and entity. The first two emphasize the fact that every term is *one*, while the third is derived from the fact that

---

<sup>227</sup> A recent example that bears out this continuing tendency may be found in Priest (2016).



every term has being, *i.e.* *is* in some sense. A man, a moment, a number, a class, a relation, a chimera, or anything else that can be mentioned, is sure to be a term; and to deny that such and such a thing is a term must always be false [omitted here by Russell: “or meaningless,” as he went on to make clear in what follows]....

Every term, to begin with, is a logical subject: it is, for example, the subject of the proposition that itself is one.... What a term is, it is, and no change can be conceived in it which would not destroy its identity and make it another term. Another mark which belongs to terms is numerical identity with themselves and numerical diversity from all other terms.... *Term* is, therefore, a useful word....

If *A* be any term that can be counted as one, it is plain that *A* is something, therefore that *A* is. “*A* is not” must always be either *false* or *meaningless*. For if *A* were nothing, it could not be said not to be; “*A is not* ” *implies that there is a term A whose being is denied, and hence that A is*. Thus unless “*A* is not” be an empty sound, it must be false—whatever *A* may be, it certainly *is*....

*Existence*, on the contrary, is the prerogative of some only amongst beings. To exist is to have a specific relation to existence—a relation, by the way, which existence itself does not have.... [W]hat does not exist must be something, or it would be *meaningless* to deny its existence....

— Bertrand Russell (1950/1903, pp. 43-44, 449-450, italics added)

In the light shed by this passage by Russell let us call our ontological question ‘the question whether non-existent things *are*’, and by formulating the question in this way, we avoid the furor that has ensued over the word ‘exist’. By employing the word ‘*are*’ in the ontological question about non-existent things we use a term that is intentionally “existentially neutral.” The word ‘*are*’ here does not embody a claim that a *certain kind* of object is in view—we are not necessarily concerned (for example) with physical objects that occupy space and persist through time, which, as we’ve noted, has become the most conventionally accepted notion of “existence.”

But the “existentially neutral” term ‘*are*’ in the question whether non-existent things *are* errs on the side of vagueness. It will come as no surprise to

readers of this book that the author recommends substituting for 'are' the expression 'are objects of reference'. This small change in wording quickly conducts us to a self-validating result that is one of the simplest and most direct applications of the de-projective method that we wish to exemplify.

Let us apply the four-stage de-projective heuristic:

Stage 1 (stipulatively descriptive): We shall stipulate that the question whether non-existent things are objects of reference means the following: By 'non-existent' we mean 'having no physical spatial-temporal dimension', and by an 'object of reference' we mean (see {8.6}) that identification/identifying reference obtains when that which is described is determined as that to which there is reference. The latter is admittedly a complex statement, but its unwieldy expression is reducible to the metatheoretical claim that when an object of reference is involved, it must be possible to determine what it is, what its identity is, relative to an appropriate reference frame. As was previously made clear ({8.6}), identification in the sense involved is very broadly understood, ranging from unique identification to general, from vague to abstract, or it may consist only in specification (again, vague or precise) according to rule.

Stage 2 (diagnostic): Let us consider two artificially and simplistically constructed, putatively competing claims, C1 and C2. C1 claims that the question whether non-existent objects are objects of reference is actually an assertion in disguise: According to C1, the "actual" disguised assertion is this: "If an object of reference *O* is not a physical object, it is impossible to refer to it." Claim C2, on the other hand, asserts that the ontological question about non-existent objects is rather this: "If an object of reference *O* is not a physical object, it *is* nonetheless an object of reference."

Stage 3 (eliminative): Consider claim C1; its meaningfulness presupposes that reference to *O* is possible. If reference to *O* is not possible, the claim "short-circuits" on the level of meaning: Assume that *O* is not a physical object; C1 claims it is impossible to refer to *O*. For C1's claim concerning *O* then to have possible meaning, the capacity of C1 to refer to *O* is presupposed. C1 denies this, and therefore C1 can have no *possible* meaning; it is an incoherent and unintelligible claim.

Claim C2, in contrast, asserts that if *O* is not a physical object, it must nevertheless qualify as an object of reference. Similar to the case of C1, were reference to *O* not possible, C2 would also be meaningless. But C2's saving grace is that instead of denying that reference to *O* is impossible, C2 affirms implicitly that such reference *is* possible. Reference to *O* is evidently possible, since the assertion of C2 involves such reference.

In short, C2 self-validates, while C1 must be rejected as projective.

Stage 4 (corrective): Out of what we've called 'respectful sympathy' for the putative meaning of C1, in an effort to express what C1 *could* mean, we find that what it really must mean—if it is to have a possible meaning at all—is what C2 itself means. Reference to *O* inescapably requires a capacity to refer to *O*.

Let us state this unashamed tautologous conclusion a little differently: The ontological question concerning the status of non-existent things must, in order to have possible meaning, metalogically entail the affirmation of the relativity of objects, of whatever kind they may *be*, to an appropriate frame of reference in terms of which they are identifiable.

### 18.3 Towards a general ontology of objects

The preceding very general and abstract conclusion reduces, when applied to specific instances, to the following family of recognitions:

- ◆ Reference to physical objects structurally/systemically presupposes an appropriate reference frame(s) in terms of which it is possible to identify such objects, as a function of which reference frame(s) they possess the identities they have.
- ◆ Reference to other sorts of objects—for example, objects of fiction—similarly presupposes an appropriate reference frame(s) in terms of which they can be identified; and so reference, for instance, to the *madeleine* in Proust's *Remembrance of Things Past* presupposes the reference frame supplied by his novel.
- ◆ Reference to abstract, theoretically based objects whose identity is specified according to rule, or is the result of a mathematical demonstration, also presupposes an appropriate reference frame(s) in terms of which these objects are identifiable. We shall accordingly find that objectivity, for example, in quantum theory ({27.13})—that is, the ontological status of quantum-theoretical objects as identifiable and re-identifiable objects of reference—is precisely framework-relative in this way, notwithstanding that the sophistication of that special variety of framework relativity is more complex.

- ◆ In general, reference to objects of whatever kind is similarly framework-relative.

Let us suppose something that today is a comparative rarity: that the reader happens to be enamored of Meinong's or Routley's theoretically refined and complex account of non-existent objects,<sup>228</sup> and let us further suppose that he or she objects to the above framework-relative conclusions. Let such a reader attempt to deny any of the above conclusions, and he or she will immediately find that metalogical projection results. For any vaguely or precisely identified object, it is projectively self-referentially inconsistent to deny that it is an object of reference.

We may state this result in an equivalent but different way: No matter of what sort, type, or level of abstraction an object of reference may be, no matter what shade or species of "existence" or of "being" or of "subsistence" that may putatively be in view, and no matter by what ontologically flavored terms that object is said to be, the recognition and affirmation of that object's framework relativity are self-validating and cannot *not* be accepted without incurring the variety of metalogical self-referential inconsistency that we've called 'projective'. Philosophers may quarrel about the words they prefer to use to characterize the different kinds of objects of reference, but such quarrels express only a taste for choosing certain words, and miss the point: *To be is to be an object of reference relative to an appropriate framework.*

This result may not be quite as sharp and brusque as Dr. Johnson's impatient kick, but it is a result with which non-philosophers—at least those who have not been bewitched by the seemingly mysterious ontology of non-existent objects—should feel quite at home. Othello, Madame Bovary, pink elephants, gluons, and the Higgs field, along with everyday physical objects, logical constructs, and the rest of the diversely populated multiverse of objects of reference—all can variously be identified, studied, and true or false statements made of them, etc., but only relative to the sometimes specialized and complex frameworks of reference which their identities presuppose.

If you wish to know the ontological status of a class of objects, look to the constitutive structure of the framework of reference in terms of which those objects are identifiable. Ingredient in any object of reference is that object's constitutive structure: i.e., the structural/systemic presuppositions of the reference frame in terms of which it is identifiable. The object, its very identity, is indissolubly linked to the reference frame(s) that provide the basis for its identifiability. The nature of that link is informative: It tells us in what sense the object *is*.

---

<sup>228</sup> E.g., Meinong (1960/1904), Routley & Routley (1973), Routley (1980).

In short, for something to *be* is a function of the coordinative relation of object of reference in relation to an appropriate reference frame in terms of which it can be identified, whether vaguely or precisely, or according to rule. Once this is fully understood, we shall see that to be an object of reference is metalogically equivalent to being an object (each metalogically entails the other), and to be a possible object is metalogically equivalent to being a possible object of reference.

## Discovery or Invention in General Problem-solving, Mathematics, and Physics

A principal objective of the epistemology of mathematics and of the epistemology of physics is to examine the claims made by mathematicians and physicists about the nature of the results they reach. One of the purposes of such an examination, a purpose that is specifically epistemological in nature, is to determine the extent to which mathematicians and physicists understand and acknowledge the conditions and limitations to which their own claims to knowledge must comply—in terms of and in relation to the frameworks of reference these disciplines presuppose and utilize in their research.

This objective of an epistemological study of mathematics and physics is self-evidently metatheoretical: The epistemologist is clearly not engaged in doing mathematics or physics, but has stepped back or above the theoretical and performative contexts in which mathematicians and physicists do their work, and in doing this he or she establishes a meta-framework of critical epistemologically focused reflection. It is not the regular task of mathematicians and physicists to engage in reflective analysis of this kind, although we sometimes, though rarely, do encounter individual mathematicians and physicists who feel a need to understand their thought and work in these terms, and, in particular, to understand the conditions and limits of the claims to knowledge which they make.

Mathematics and physics are exemplars of disciplines that historically have been the most successful in reaching results in a progressive, incremental manner, results to which the efforts of previous generations of mathematicians and physicists have contributed and have made possible the evolution and advancement of knowledge. In addition, the variety of knowledge that is progressively acquired by both mathematics and physics also stands as a prototype of knowledge that is obtained using methods that are among the most rigorous, orderly, and systematic. These three properties, when successfully combined, have provided humanity with methods that (i) are *precise*, (ii) impose upon participating researchers prerequisite *shared standards* of what such orderly research is stipulated to mean, and (iii) are *systematic* in the fundamental sense of this word, which refers both to the

methodological rule-based definitions of the procedures that must be followed by mathematicians and physicists, and to the general-systems community of researchers that is formed when groups of specialists adhere to and implement a shared set of methodological and disciplinary rules.

Understood in this way, both mathematics and physics comprise very nearly ideal prototypes of models of problem-solving. Before moving on in this chapter to consider each discipline separately in an applied epistemological context, it will be useful first to consider them together in the context of closely related epistemological issues that arise in connection with general problem-solving.

### 19.1 Mathematics and physics from the standpoint of the general theory of problem-solving<sup>229</sup>

In its most abstract sense, a *problem* defines a gap between a set of initial conditions and a desired goal. Just as the epistemology of mathematics and the epistemology of physics involve a critical, reflective step back from or above the routinely employed frameworks of the two disciplines, so does the epistemology of general problem-solving. Here, too, a major purpose of an epistemological examination of general problem-solving is to determine the extent to which problem-solvers understand and acknowledge the conditions and limitations with which their own results must comply—in terms of and in relation to the frameworks of reference which they presuppose and utilize in their efforts to solve problems.

Theorists who have developed an interest in studying solvers of problems have sought to understand the analytical and intuitive processes involved in problem-solving. They have recognized that problems may be of a great many kinds and possess many different degrees of complexity, and that problems vary greatly according to the ways in which they are formulated by researchers in any given discipline. A large number of these theorists have been psychologists, some have been philosophers, some have been researchers in artificial intelligence, and some have come from other disciplines. In their efforts to understand the nature of problem-solving, they have come to make numerous observations about the nature of problem-solving, and—a subject that will concern us here in particular—about the nature of the relationship between solutions to problems and the problems which they solve.

The relationship between a problem and a solution to that problem has been interpreted by problem-solving theorists in two opposing ways. Their

---

<sup>229</sup> This section is based in part on Bartlett (1978a).

disagreement about this relationship has been a very fundamental disagreement, so much so that the position taken by a theorist of general problem-solving, as well as the position taken by the individual problem-solver, often inclines or predisposes both the theorist and the problem-solver to favor certain directions of research rather than others. This basic disagreement is, in other words, not only theoretical, but it has practical consequences in terms of how problems may actually be approached and solved. As we shall see in this chapter, this disagreement is epistemologically problematic in ways that the metalogic of reference can resolve.

The disagreement has to do with whether solutions to problems are to be considered “*discoveries*,” or whether solutions instead comprise the “*inventions*” or “*creations*” of the problem-solver. The term ‘*discovery*’ of course has a number of different meanings; the meaning that relates to problem-solving is what we commonly think of as involving an action or a process of finding out or of becoming conscious of something *for the first time*, of something which was *previously* not known, but which nonetheless *previously*, in some sense, *was*. Associated with this meaning is the notion that an action or process which leads to a discovery is one which *discloses* or *brings to light* something that was previously *hidden* from sight or simply *inaccessible*, whether physically or cognitively. A genuine discovery in this sense reveals something that is *found for the first time*, and yet that something is nonetheless commonly believed to have been the case, or to have existed, *prior* to its discovery. This meaning of ‘*discovery*’ and of ‘*discovering*’ derives from the Anglo-Norman ‘*discoverir*’, meaning “to uncover.” The word was often used with the implication that a physical obstacle or covering is removed, bringing that which was previously unknown out of concealment.

The word ‘*invention*’ and the verb ‘*to invent*’ are derived from the Latin ‘*invenire*’, meaning literally “to come upon.” To invent something has acquired the meaning of *fabricating*, or *making something up*, perhaps something that is a fabrication in the sense of what is *fictitious* or *false*. Added to this is the belief that what is “made up” is created *for the first time*; it is constructed through originality; its *origin* is to be found in the *act of creation*; what is invented *did not exist before*; it is *new* and was *previously unknown*. With this collection of meanings, we say of a creative work of art or literature that it was “composed”—i.e., it was brought into being by the artist or writer; in no sense did it exist before; it was through a creative, inventive process that the work came to be.

In short, what most distinguishes “discoveries” from “inventions” is the claim that what is discovered or invented did or did not exist beforehand. As we shall see, it is precisely this claim that leads to epistemological problems.



Among early researchers who have embraced the discovery model of problem-solving are Bruner (1961), Pólya (1962-1965), and Skinner (1966, pp. 235, 247); among early problem-solving theorists who have promoted the invention model are Guilford (1958), Taylor (1958, 1964), Getzels & Jackson (1962), and Gruber, Terrell, and Wertheimer (1962). Their difference of opinion has been expressed in conflicting answers they have given to these questions: Is the solution of a problem generally a matter of human discovery, or of human creativity? Is the solution of a problem autonomous of human problem-solving activity, or is the solution the result of human faculties and therefore a creative invention?

When a problem is solved, both problem-solving theorists as well as problem-solvers themselves frequently believe that the solution that has been *found* constitutes a “discovery” when there is an evident sense in which the solution “was already there.” In this sense, Columbus is said to have “discovered” America because there is evidence that America *was there before* he arrived on its shores. In a related but different sense, a problem-solver is said to have “discovered” the solution to a set of three simultaneous equations in three unknowns because the values of the unknowns were already given implicitly by the equations, and in that sense could be said “already to have been there.”

To call the solution to a problem a result of the “creative” faculties of the problem-solver is to make a contrary claim, that the solution “was not already there,” but rather was obtained through the “inventive” resources of the problem-solver.

As was mentioned earlier, the way in which problem-solving is interpreted can have significant and practical consequences: On the one hand, if a problem-solver conceives of problem-solving in terms of the discovery model, he or she may be likely to attempt to formulate objective heuristic principles, formal techniques, and guidelines that can furnish problem-solvers with tools or methods that can enable them to reach solutions that are “already there,” waiting to be disclosed. On the other hand, if the problem-solver interprets problem-solving as essentially a creative process, he or she may be likely to focus on the psychology of invention in an attempt to gain some understanding of useful characteristics common to creative problem-solvers. As Gagné (1966, p. 129) observed many years ago, if problem-solving were in fact a matter of discovery, consisting perhaps of no more than the application of relevant objectifiable rules and methods to reach solutions, such problem-solving would likely be thought *too routine* to qualify as genuine problem-solving from the standpoint of problem-solvers who embrace the creative model. But, from the opposing point of view, for the discovery-

oriented problem-solver, “creative processes” are likely to be thought *too vague* to provide a systematic basis for an adequate and effective approach to problem-solving. An approach to problem-solving that considers the application of objectifiable rules to provide the best route to the discovery of solutions can lead directly to problem-solving by means, for example, of algorithmic programs or artificial intelligence. The opposing approach that centers attention on the role of the creative problem-solving subject will emphasize the role of individual psychology.

The question is, of course, whether either view can be justified. The two views come into conflict in connection with the alleged status of the solution to a problem *before* the subject solves the problem. If there is some justifiable sense in which the solution “was already there,” then support can be given for the discovery model. If justification can be provided for the view that the solution “was not already there,” then in this sense it is commonly said that the problem-solver was creative.

Two things are immediately evident, but they are also trivial: For any problem-solver a problem expresses his or her initial ignorance of a solution. In this obvious sense, the solution “was not already there” for the would-be problem-solver before he or she solves the problem. However, for other problem-solvers who knew the solution beforehand, the problem has a known solution; that solution “was already there” for those who knew the solution beforehand. In these trivial senses, every problem is solved “creatively,” and some of the solutions are “discovered.”

The question becomes more interesting in connection with the solution of a problem not known to have been solved before. Is such a solution a “discovery” or an “invention”? Can justification be given for either claim?

But with such a problem, in neither case is justification *possible*. We recognize that to determine whether a solution is “discovered” or “invented” can only be settled by reference to the status of the solution before the problem was solved. In the case of a problem with no known solution, did the solution exist, implicitly or explicitly, in any sense before the problem was solved for the first time? Given that solving a problem provides us with knowledge of its solution, then prior to solving the problem no knowledge concerning the solution is possible. It then follows that prior to the solution of a problem no knowledge regarding the “existence of the solution, implicitly or otherwise” is possible.

In this sense—and this is the non-trivial sense of the conflict between the two views in question—any attempt to talk about conditions of either “discovery” as opposed to “invention” fails to satisfy the preconditions of possible reference to the status of a solution prior to the solution of the problem. The

conclusion is uncompromising: When a previously unsolved problem is involved, in neither case is it possible meaningfully to claim that problem-solving is a matter of discovery or of invention. To make either claim is to trespass beyond a problem's horizon of possible meaningfulness.

### 19.2 Discovery or invention in mathematics

- (a) Assume that a calculation is made which gives a result  $R$ . We double-check the calculation, confirm its validity, and prove that  $\sim R$  leads to contradiction. Hence  $R$  was true before the calculation was made.
- (b) Suppose that on the future date of Jan. 1, \_\_\_\_\_, it is proved that an infinity of twin primes exists. Let us call the proof of this result ' $P$ '. Therefore,  $P$  was true before that date.
- (c) An epistemological analysis of the presuppositional structure of a particular knowledge claim in mathematics leads to a result  $M$ . We find that  $\sim M$  leads to metalogical self-referential inconsistency. Therefore,  $M$  must be true, prior to and independently of this epistemological analysis.

The conflict between the discovery and invention models that I have described in connection with general problem-solving is closely paralleled by a similar conflict between the realist-logicist and the intuitionist conceptions of mathematics. These opposing views have been remarkably tenacious, the first having been held more than two thousand years ago by Plato and in the last century by Frege and Gödel, and the second view was in a sense supported by Kant in the context of his own special conception, and more recently was developed by intuitionists L. E. J. Brouwer and Arend Heyting.

Frege expressed one side of this conflict in these words:

[E]ven if all reasonable creatures should at some time simultaneously slip into hibernation, the truth of [a true mathematical] ... statement would not, as it were, be suspended for the duration of this sleep, but would remain undisturbed. The truth of a statement is not its being thought....

[T]he matter is elevated from the realm of the subjectively possible to that of the objectively definite. Indeed, the fact that from certain statements another statement follows is

something objective, something independent of whatever laws may govern the wanderings of our attention; and it makes no difference whether we really make the inference or not. (Frege, 1964/1884, pp. 101, 103)

Intuitionist Arend Heyting did not agree:

[M]athematics is a production of the human mind.... [W]e do not attribute an existence independent of our thought, i.e., a transcendental<sup>230</sup> existence, to the integers or to any other mathematical objects.... Even if they should be independent of individual acts of thought, *mathematical objects are by their very nature dependent on human thought. Their existence is guaranteed only insofar as they can be determined by thought. They have properties only insofar as these can be discerned in them by thought....* Faith in transcendental existence, unsupported by concepts, must be rejected as a means of mathematical proof....

A mathematical proposition expresses a certain expectation. For example, the proposition ‘Euler’s constant  $C$  is rational’, expresses the expectation that we could find two integers  $a$  and  $b$  such that  $C = a/b$ .... The ... [expectation] refers not only to a state of affairs *thought to exist independently of us* but also to an experience thought to be possible.... (Heyting, 1964, pp. 42, 47, italics added)

In the above passage we immediately notice the incongruity between Heyting’s claim that “mathematical objects are by their very nature *dependent* on human thought,” and his claim that a mathematical proposition “refers ... to a state of affairs thought to exist *independently* of us....” We shall return to this incongruity later.

For realist/logicist mathematicians like Plato and Frege, it felt natural to judge mathematical results to be “discoveries,” while intuitionist mathematicians like Heyting and Brouwer claim that such results are the consequences of the creative activity of the mathematician’s mind, and are, in this sense, “inventions.” For the realist/logicist, the solution to a mathematical problem is claimed to be autonomous of human industry; for the intuitionist, it is claimed to be essentially reliant upon human faculties.

The realist/logicist mathematician would agree with the first two claims

---

<sup>230</sup> Heyting meant (or should have used) ‘transcendent’ here and subsequently.

posed at the beginning of this section, and would probably also affirm the third. The intuitionist would likely deny all three assertions. Let us look at these assertions one at a time:

The claim (a) was: Assume that a calculation is made which gives a result  $R$ . We double-check the calculation, confirm its validity, and prove that  $\sim R$  leads to contradiction. Hence  $R$  was true before the calculation was made. — As we reflect on this claim, we need to ask in what sense, if any, is a frame of reference *possible* in terms of which the contradiction-entailing  $\sim R$ , and hence the inferred necessity of affirming  $R$ , can be recognized prior to the calculation in view? Clearly, if another or different calculation had already previously established the validity of  $R$  by showing that its negation leads to contradiction, then raising this question would be self-answering and pointless. We therefore stipulate that the calculation at issue is the first calculation ever to have shown that  $\sim R$  leads to contradiction.

When this is the case, it would be projective to give the result  $R$  retrospective validity, since the reference frame necessary for its demonstration would, under the above conditions, fail to be possible.

Claim (b) at the beginning of this section was this: Suppose that on the future date of Jan. 1, \_\_\_\_\_, it is proved that an infinity of twin primes exists. (A prime number has a “twin” when that twin is 2 more or less than it—e.g., 41 and 43 are twin primes.) Call the proof of this result ‘ $P$ ’. Therefore, according to claim (b),  $P$  was true before that date. We now wish to answer the question, Was  $P$  true before that date or not?

The problem whether there is an infinity of twin primes is known as the twin prime conjecture, which—as of the present writing and despite numerous attempts to solve the problem and some potentially significant progress—has not yet been proved.

In a paper written in dialogue form, Heyting has his character, The Intuitionist, respond to the above question in the following terms:

A mathematical assertion affirms the fact that a certain mathematical construction has been effected. It is clear that before the construction was made, it had not been made. Applying this remark ..., we see that before Jan. 1, [ \_\_\_\_\_ ] [the assertion] had not been proved....<sup>231</sup> But this is not what you mean. It seems to me that in order to clarify the sense of your question you must ... refer to metaphysical concepts: to some world of mathematical things existing independently of our

---

<sup>231</sup> The date which Heyting employed in the blank was 1970. He published his paper 1956, copying the then-future year that was originally used in this example by Menger (1930).

knowledge, where [the proposition] is true in some absolute sense. But ... mathematics ought not to depend upon such notions as these. In fact all mathematicians and even intuitionists are convinced that in some sense mathematics bear upon eternal truths, but when trying to define precisely this sense, one gets entangled in a maze of metaphysical difficulties. The only way to avoid them is to banish them from mathematics. This is what I meant by saying that we study mathematical constructions as such.... (Heyting, 1964/1936, pp. 56-57)

Assertion (b), that  $P$  is true before it was proved to be true, takes a famous conjecture in mathematics but raises a question that is fundamentally the same one, but in different clothing, as we already dealt with in connection with assertion (a). The claim made by (b) concerns the judgment that we should be prepared to make about a future contingency, the possible proof of  $P$ , but other than this future-focused contingency, assertions (a) and (b) both *attempt* to ascribe validity retroactively in some—putatively—meaningful sense to a result that can only be reached by means of the reference frame of a proof that is assumed not to have existed, (a), or not yet to exist, (b). This attempt expresses the motivation of those who embrace the discovery model in mathematics; it is an attempt to claim that mathematical results exist in some sense that is autonomous of the mental processes of individual mathematicians.

Heyting's choice was "to banish" such assertions as are involved in discovery claims, not wishing, as he expressed this, to become "entangled in a maze of metaphysical difficulties." But doing this appears to reflect no more than a *preference*, a taste—certainly one that is deeply seated and deeply felt by intuitionist mathematicians, but still it remains a preference unless it can be proved. A proof that "metaphysical mazes are to be avoided" is not to be found in intuitionist mathematics. However, our metalogical reasoning in the present study does not express a mere preference in favor of banishing metaphysical claims, but instead we've come to a point in this study at which we can easily recognize that the claim made by the discovery theorist in mathematics cannot, in principle, *possibly* be made meaningfully because to attempt to make that claim is to fall victim to self-undermining projection: It is an attempt both to make a claim for which a certain reference frame is indispensable, and yet, at the same time, to deny that such a reference frame is available. It therefore can make no sense either to claim that  $P$  was or was not true before January 1, \_\_\_\_\_. We may only claim that the truth of  $P$  is a function of the framework enabling its proof on January 1, \_\_\_\_\_. We shall call projections of this kind '*projections of discovery*'.

What of the invention model in mathematics? L. E. J. Brouwer ascribes mathematical existence to the intuitively based creative faculties of mathematicians. He sought to analyze the way in which mathematical thinking proceeds. For Brouwer, it is essentially a temporally based intellectual process in which what he calls “the intuition of two-oneness” plays a fundamental cognitive role. In his view, this intuition refers to the temporally successive nature of human consciousness that underlies the very experience of counting and the concept of number. The following is a typical passage that we find in Brouwer’s writing:

[T]he falling apart of moments of life into qualitatively different parts, to be reunited only while remaining separated by time, [is] the fundamental phenomenon of the human intellect, passing by abstracting from its emotional content into the fundamental phenomenon of mathematical thinking, the intuition of bare two-oneness. This intuition of two-oneness, the basal intuition of mathematics, *creates* not only the numbers one and two, but also all finite ordinal numbers, inasmuch as one of the elements of the two-oneness may be thought of as a new two-oneness, which process may be repeated indefinitely.... (Brouwer, 1964/1912, p. 69)

In a later paper, Brouwer made his epistemological commitment more explicit:

[T]hat *there are no non-experienced truths* ... has found acceptance with regard to mathematics.... Mathematics rigorously treated from this point of view, and deducing theorems exclusively by means of introspective construction, is called intuitionist mathematics.... (Brouwer, 1964/1940, p. 78, italics added)

The status of the principle of excluded middle (Brouwer calls it ‘the principle of the excluded third’) has been a source of debate for millennia among philosophers and logicians. Referring to the principle, Brouwer claimed:

For intuitionism the principle of the excluded third and its corollaries are assertions  $\sigma$  about assertions  $\tau$ , and these assertions  $\sigma$  only then are “realized”, i.e., only then convey

truths, *if these truths have been experienced.*” (Brouwer, 1964/1940, p. 80, italics added)

We therefore find that very explicit in Brouwer’s intuitionist epistemology is the claim that mathematical results are created through the thought processes of individual intuitionist mathematicians, i.e., the very contrary of the claim made by the logicist/realist. To assert, as Brouwer did, that “there are no non-experienced truths” is to maintain that mathematical results are *not autonomous* of their construction by the minds of mathematicians. And just like the conflicting claim made by the logicist/realist, it is, and for the same metalogical reasons, projectively self-undermining. To claim that mathematical results *depend* upon the thoughts of mathematicians is to claim that, in some putatively meaningful sense, these results would no longer—what shall we say?—“be valid” or “exist” were a mathematician’s mind not employed to prove those results.

Here again, given the statement of the conditions that define the intuitionist point of view, there is no *possible* reference frame in terms of which such a dependency claim can meaningfully be made. There is no possible frame of reference from the standpoint of which to have access to any mathematical result without already presupposing the reference frame in terms of which its proof is possible. We shall call projections of this kind ‘*projections of invention*’.

When we consider the opposing positions of the logicist/realist and the intuitionist and find that neither position which each has taken with regard to the status of mathematical results can, in principle, be meaningful, we reach another vantage point from which to appreciate the pervading nature of framework relativity.

As Kant warned long ago (see above, {14}), there is a very human tendency to “overstep the boundaries” of a given domain of intellectual effort, and to “go astray in delusion....” Both logicist/realist mathematicians and intuitionist mathematicians succeed in reaching what they judge to be mathematical truths; and yet, each—again to use Kant’s words—feels “compelled, notwithstanding all interdictions against losing himself in transcendent ideas, to seek rest and contentment beyond all the concepts which he can vindicate by experience...” (Kant & Beck, 1950/1783, §57).

One of the comparatively rare mathematicians who sometimes struggled successfully against this compulsion was David Hilbert. Although the following excerpt is not altogether coherent in terms of the epistemological framework of the metalogic of reference, here and there it comes close. In these passages, Hilbert discusses the concept of the infinite in mathematics:



[T]he infinite is nowhere to be found in reality, no matter what experiences, observations, and knowledge are appealed to.... Does material logical deduction somehow deceive us or leave us in the lurch when we apply it to real things or events?... It deceives us only when we form arbitrary abstract definitions, especially those which involve infinitely many objects. In such cases we have illegitimately used material logical deduction; i.e., *we have not paid sufficient attention to the preconditions necessary for its valid use.*

*As a ... precondition for using logical deduction and carrying out logical operations, something must be given in conception....* [T]he subject matter of mathematics is, in accordance with this theory, the concrete symbols themselves whose structure is immediately recognizable....

[With regard to infinite disjunctions:] Such an extension into the infinite is, unless further explanation and precautions are forthcoming, no more permissible than the extension from finite to infinite products in calculus. Such extensions, according to usually *lapse into meaninglessness....*

Our principal result is that the infinite is nowhere to be found in reality. It neither exists in nature nor provides a legitimate basis for rational thought—a remarkable harmony between being and thought.... Operating with the infinite can be made certain only by the finitary.

The role that remains for the infinite to play is solely that of an idea— ... an idea which we may unhesitatingly trust *within the framework erected by our theory.* (Hilbert, 1964/1925, pp. 142, 144, 151, italics added)

Despite his implicit recognition of the framework relativity of his mathematics, in the above passage Hilbert transgressed the horizon established by his own framework: His claim that the infinite is *not found in reality* is projective. Given his own framework-relative position, there is no possible reference frame from the standpoint of which the truth of this claim could be determined.

Unlike logicist/realist and intuitionist mathematicians, Hilbert was a formalist: As he expressed this, “the subject matter of mathematics is, in accordance with [my] theory, the concrete symbols themselves whose structure is immediately recognizable” (Hilbert, 1964/1925, p. 142). Here we find, once

again, the expression of a *preference*: There is no possible reference frame in terms of which, for example, Hilbert could demonstrate that the subject of mathematics is purely and only a formalism, and (i) *not* what the intuitionist claims it is, i.e., essentially a matter of human thinking, and (ii) *not* what the realist/logicist claims it is, i.e., a matter of autonomous mathematical reality.

Let us conclude this section by extending our results to the third assertion (c) quoted at the beginning. It was: An epistemological analysis of the presuppositional structure of a particular knowledge claim in mathematics leads to a result  $M$ . We find that  $\sim M$  leads to metalogical self-referential inconsistency. Therefore,  $M$  must be true, prior to and independently of this epistemological analysis.

We recall that in connection with the problem of putative meaning ( $\{12\}$ ), the framework relativity of metalogical analysis itself was recognized. The results reached by the metalogic of reference are themselves accessible, potentially true, and meaningful *as a function of* the metalogical framework this study formulates. To ask whether such results are true, or whether they fail to be true, *prior to* and *independently* of that framework is to ask a question that, in principle, is without possible meaning. "True" in the context of (c) means "true relative to metalogical analysis." Similarly it without possible meaning to claim that the truth of any true results that are reached are "created" by the metalogical frame of reference presupposed to demonstrate them. The conditions that would have to be satisfied to determine that the results reached were non-existent, or not the case, prior to metalogical analysis (and hence were "created"), cannot be met without presupposing the reference frame of the analysis itself.

The human temptation is commonly to affirm the negation of a given position when that position is shown in some way to be self-defeating or self-contradictory. Yet, we need to discipline ourselves to resist generalizing this tendency, for the negation of a projective theory, claim, or concept retains the projection which is to be rejected; the negation of a projection is, as expressed in  $\{11.4\}$ , "infected" by such meaninglessness.

Once one has attempted to transgress beyond the boundaries of possible meaning, affirmations and denials beyond those boundaries can themselves make no possible sense. We shall see this repeatedly, but in different guises, in the chapters that follow. We shall encounter a variety of expressions of the projective compulsion, which is natural, normal, and habitual for most people, including many scientists and philosophers.

### 19.3 Discovery or invention in physics

Newton's laws, the principle of contradiction, any truth whatever—these are true only as long as [man] *is*. Before there was [man], there was no truth; nor will there be any after [man] is no more. For in such a case truth ... *cannot* be. Before Newton's laws were discovered, they were not "true"; it does not follow that they were false, or even that they would become false if ... no [such discoveries] were any longer possible....

To say that before Newton his laws were neither true nor false, cannot signify that before him there were no such entities as have been uncovered and pointed out by those laws. Through Newton the laws became true; and with them, entities became accessible in themselves to [man]. Once [these] entities have been uncovered, they show themselves precisely as entities which beforehand already were....

– Martin Heidegger (1962/1927, H226-227)<sup>232</sup>

Heidegger, to be sure, was no physicist and did not pretend to be. I quote this passage, not because it is epistemologically sophisticated, which it is not, but because, due to its very lack of critical analysis, it focuses attention on a group of epistemological problems relating to the question whether physical laws and principles are "discovered" or "invented" in physics. Heidegger's impressionistic and typically mystifying passage, written in 1927, was a philosophical precursor of epistemological controversies, as we shall see in this and later chapters, which were to take place among leading physicists as well as among philosophers of science during the subsequent golden decades of general relativity, a period accompanied by pioneering developments in quantum theory.

If we make an effort to translate Heidegger's passage into the language of this study, he asserts a group of claims relating to physics, among them: that physical laws are fundamentally framework-relative; that their truth cannot be dissociated from the conceptual framework in terms of which they are formulated; that, apart from (or prior to the establishment of) a particular physical theory, such physical laws are (or were) neither true nor false; that those

---

<sup>232</sup> For readers unaccustomed to Heidegger's often mystifying language, I've made a few editorial changes in this passage in order to provide some clarity and no loss of meaning given our purposes here. I've substituted 'man' for Heidegger's '*Dasein*' and edited out other similar phraseology.

laws that are judged to be “true” of the physical world are “discovered”; and that such laws “become truths” when a physicist, such as Newton, “discovers” them. Heidegger then proceeds to refer to what he calls ‘entities’—and here let us make the shorthand assumption that this includes physical phenomena such as events, physical objects, waves or fields, etc. And then he adds an interesting claim: that once such physical laws are “discovered,” the “entities” which these laws refer to “present themselves” (i.e., are experienced) as “entities which beforehand already were.” In other words, physical phenomena that are conceptualized by a given physical theory are, according to Heidegger, experienced as comprising phenomena which “already were” before that theory made it possible to recognize them.

Since I do not intend to digress in a critical commentary on Heidegger’s thought concerning physical laws, let us suppose that the above interpretive translation is acceptable as it stands: Our interest will not be in Heidegger’s thought, but in the content of the above translation which contains in miniature a cluster of assertions made by a number of leading physicists when they have sought to give an account of their epistemological positions.

In the previous section of this chapter, the conclusion was reached that both discovery theorists and invention theorists in mathematics fall victims to projective incoherence when they affirm that mathematical results exist independently of, or are dependent upon, the thought processes of mathematicians. It is one thing to reach a conclusion like this when a domain of wholly abstract objects of mathematics is in view, and quite another to extend the same recognition beyond the formal world of mathematics to the common world of the great majority of people who, along with many philosophers, believe in an “independently existing,” “external” world that, for most physicists, comprises the focus of their study.

Let us turn again to the passage quoted from Heidegger: If one accepts the consequences that follow from the present study’s understanding of framework relativity, then it *can make no sense* even to ask whether there are truths (or falsehoods) apart from the presupposed framework capable of answering this question. Although Newton’s laws, in relation to the physical phenomena to which these laws referred when Newton formulated his laws, *were* at one time judged universally to be true of the physical universe, they are no longer universally accepted; in particular, they have been superseded and replaced by relativity physics when, for example, high velocities, acceleration, and gravitational forces enter in.<sup>233</sup> It is projectively meaningless to refer to physical

---

<sup>233</sup> Newtonian mechanics remains useful in a great many applications, such as calculating bridge strengths or orbital trajectories. The point made in the text relates to the fact that Newton’s laws are no longer universally applicable.

phenomena—as governed *universally* by the laws which Newton’s framework established—either *before* that framework was available, or *after* that framework was superseded. Heidegger, in my interpretation, suggests that apart from (or prior to the establishment of) a particular physical theory, such physical laws are (or were) neither true nor false. However, it is not a matter of these laws being neither true nor false under such conditions, they rather are devoid of possible meaning.

In order to avoid projectively self-undermining claims, this recognition compels us to accept the simple fact that what physicists understand by the physical universe is no longer what it was understood to be (for example, at the time of Newton): In terms of the present understanding of “the physical universe,” *that which we understand is different*. For example, for physicists who have renounced the physically meaningful separability of the observed and the observer, relative to their currently accepted theoretical framework, a different law-governed, post-Newtonian physical universe is now the object of study.

Such a recognition is the basis for the metalogical claim that it is impossible meaningfully to refer to the (or a) physical universe in a way that is framework-independent, and hence that our *understanding* of the universe and the universe *as we can meaningfully refer to it, cannot but be one and the same*. This sense of “understanding” the physical universe is, in other words, a function of the conceptual framework presupposed by reference to it.

The final phrase in the quoted passage from Heidegger is clearly problematic: He claimed that once a given physical theory conceptualizes that which it takes to comprise physical reality, and a group of physical laws is “discovered” to hold true of that physical reality, the physical phenomena (the “entities”) which these laws refer to are experienced as phenomena that “already were” *before* the theory was formulated in terms of which it was possible to recognize them. Whether Heidegger intended the word ‘before’ to have a temporal or a logical meaning is not crucial to us here, for in either case we have a stereotypical case of projective meaninglessness. In either meaning of the retrodictive word ‘before’, Heidegger pulls the carpet out from under his own feet: We find here yet another example of the human compulsion to trespass beyond the metalogical boundaries of a given reference frame. It simply cannot—in principle—be meaningful to claim that, *apart* from the reference frame established by a given physical theory (or established by another equivalent theory isomorphic with it), the set of physical principles it establishes either is true, or is false, of physical reality.

What should we then make of the human experience reported by Heidegger, the experience of believing that physical laws that have been found to be

true, along with the phenomena which they describe, were somehow “already there” and were therefore “discovered”? When we reflect on the putative meaning of this experience, we find that it *can be* nothing more than the expression of the widespread human psychological and cognitive disposition to fall victim to a form of delusional thinking that interprets phenomena as though they have putatively meaningful autonomy in relation to the reference frame(s) that makes reference to them possible.

As is typical of projective claims, the contrary assertion, that physical laws are “invented,” is projective, for the same reason.

These considerations compel one to acknowledge that the concepts of “discovery” and “invention” when applied to the results reached by physicists are evidently of a different kind than we encountered in mathematics. Physical results differ from mathematical results in that they are not solely a function of (determined by and derived in terms of) a formalized system or formal thought process, but physical results are more than this: When they are verified to hold true, a coordinative relationship<sup>234</sup> is established between the formalized language or formal thought process employed by the physicist and the set of phenomena that defines what the physicist conventionally understands by “the real world.” This relationship was boldly and—as we shall see in detail later ({27.4})—projectively expressed in the opening paragraph of Einstein’s, Podolsky’s, and Rosen’s now-famous paper of 1935:

Any serious consideration of a physical theory must take into account the distinction between the objective reality, which is independent of any theory, and the physical concepts with which the theory operates. These concepts are intended to correspond with the objective reality, and by means of these concepts we picture this reality to ourselves. (Einstein, Podolsky, Rosen, 1935, p. 777)

As we shall discuss in later chapters ({26, 27}), this passage expresses a belief that is no longer embraced by all physicists. We perceive in a passage like this that the kind of “truth” which physicists like Einstein, Podolsky, and Rosen believe they formulate involves a relationship between a set of phenomena that is considered to be *other than* and *independent of* their own theoretical representation. Just what this ‘other than’ and ‘independent of’ mean has long been a source of debate and controversy among physicists and philosophers of science. The status accorded to such framework-independence determines whatever meaning might, in principle, be associated with the con-

---

<sup>234</sup> See {5.7}.

cepts of “invention” and “discovery” when these are applied to physical laws. But the conclusion that is forced upon us is that it is in principle not possible to provide a basis for such meaning.

#### 19.4 Discovery or invention, according to Einstein

Einstein used the terms ‘discovery’ and ‘invention’ as they apply to theoretical physics in several of his publications. One finds that in most of his published work Einstein claimed that the principles and laws formulated by theoretical physics are not inferred from experience, but are what he called ‘free inventions’ of the human mind. He made this claim frequently.<sup>235</sup> The following passage is characteristic:

The concepts which arise in our thought and in our linguistic expressions are all—when viewed logically—the free creations of thought which can not inductively be gained from sense-experiences. (Einstein, 1944, p. 287)

Not only did Einstein believe that the fundamental principles of physics are the products of the “free invention” of the thought of the physicist, but much like intuitionist mathematicians, he similarly believed that “the series of integers [is] an invention of the human mind” (Einstein, 1944, p. 287). The reader should by now be aware of the projective nature of such a claim.

And yet, like many physicists as well as non-physicists, Einstein also was attracted by the notion that physical laws are “discovered”:

I am convinced that we can *discover* by means of purely mathematical constructions the concepts and the laws connecting them with each other, which furnish the key to the understanding of natural phenomena...” (Einstein, 1934, p. 36, italics added).

And then, on the next page, he continued by saying: “In a certain sense ... I hold it to be true that *pure thought* can grasp reality, as the ancients dreamed” (Einstein, 1934, p. 37, italics added).

If one reviews his publications, we find that Einstein was comfortable using both terms, ‘discovery’ as well as ‘invention’: Theoretical physics, in his view, involves a conceptual construction in the mind of the physicist, a

---

<sup>235</sup> E.g., Einstein (1933, 1936, 1944).

“free invention,” which is then, if successful, “discovered” to “correspond with the objective reality” (as he, Podolsky, and Rosen expressed this).

One might innocently be led to think that the epistemological question, whether the principles formulated by theoretical physics are “discovered” or “invented,” could be laid to rest in much the same way as was the similar question in mathematics. But in physics we find that a purported relationship is ascribed, a relationship that is believed to obtain between objects belonging to two very different categories: mind and physical reality. Physicists like Einstein, Podolsky, and Rosen believe that this relationship putatively obtains between the conceptual constructions of the physicist and phenomena that are “other than and independent of” those constructions. —Here, we find ourselves face-to-face with the famous and epistemologically central “problem of the external world,” which we shall examine in detail in {21}.

Let us sum up what it is possible, in principle, meaningfully to claim about the status of physical laws—in particular, whether they are either “discovered” or “invented.” The conclusion we have reached in this chapter is that we face a metalogically limiting situation when a recognition of the solution to a mathematical, physical, or, in fact, any general problem occurs *for the first time*. When the solution to such a problem is recognized for the first time, the conditions presupposed by that recognition are such that it is projectively meaningless to claim that either the solution was “discovered” or “invented.” Instead, the de-projective position we reach affirms the fact that a recognition of the solution to a problem brings with it the necessity of associating it with the frame of reference (or an equivalent frame) that was involved in its apprehension, and that to attempt to transgress beyond the metalogical horizon established by that reference frame is to fall victim to meaninglessness. Claims about the status of mathematical theorems or about the status of physical principles cannot meaningfully be made except in essential relativity to the frameworks of reference in terms of which they are *capable* of being formulated and confirmed.

An epistemological result of this kind has surprisingly far-reaching consequences, not only for philosophers, but perhaps also for physicists, as we’ll see in later chapters. We might recall that Einstein, who was also intensely concerned with the epistemological basis of his work, wrote:

It may appear as if all such considerations were just superfluous learned hairsplitting, which have nothing to do with physics proper. However, it depends precisely upon such considerations in which direction one believes one must look for the future conceptual basis of physics. (Einstein, 1959/1949a, p. 683)



## 20

### The Conceptually Unreachable: “*The Far Side*”

[T]here are ... facts that are unknowable in principle—for anyone and everyone as a matter of inexorable necessity.... [T]he questions that will preoccupy [me] are not just questions we cannot answer, but questions which, in the very nature of things, no one can possibly answer.

– Nicholas Rescher (2009, p. 2)

[T]here is something inherently contradictory in the very project of theorising about limits of thought. In the very process, one is required to conceive or describe things that are on the other side....

Of course, one might reject the contradiction by rejecting the claim that there are things beyond the limits of thought.

– Graham Priest (2007, p. 177)

This chapter examines a set of claims that philosophers have made concerning what we might metaphorically call ‘*The Far Side*’. Assertions of this variety which have tantalized philosophers and which I relegate to “The Far Side” include such claims as these: There are truths and facts that we cannot, in principle, know. That is, there exist truths and facts that we not only *do not* know, but *cannot* know since they exceed any entity’s *possible capacity* to know them. Alternatively, some philosophers have expressed an interest in questions which, they claim, we not only have no answers for, but which are in principle unanswerable. These, so they claim, are questions that no entity *can possibly* answer. Still other philosophers have claimed that we can think thoughts beyond the limits of our own thinking.

We shall call such alleged truths, facts, questions, and thoughts ‘*conceptually unreachable*’.

Claims about “the conceptually unreachable” define the domain of “The Far Side.” They are claims that relate specifically to what is asserted by some philosophers, including some logicians, to lie beyond possible capacities of knowledge—not only beyond human capacities, but beyond the possible abilities of any knower. They are claims about certain questions which, some have argued, not only have not been answered, but can never, ever, in principle, be answered, by anyone. And among these claims about “the conceptually unreachable” is the occasional claim that it is possible for thought to reach beyond its own limits.

These assertions that concern “The Far Side” relate—*putatively*, and this needs to be emphasized—to truths, facts, answers to questions, and thinkability which, as we shall see, are believed to lie, in various ways, beyond the metalogical horizon of any possible frame of reference.

### 20.1 “The conceptually unreachable” and the concept of horizon

To be sure, one who believes a proposition without believing its more obvious consequences is a fool; but it is an empirical fact that there are fools.

– Alonzo Church (2009, p. 14)

The subject of “the conceptually unreachable” should immediately bring to the reader’s mind the development in this study of the concept of metalogical horizon. As we have seen in previous chapters, a frame of reference is generally subject to a group of referential limitations, among them:

- ◆ It has *limits*: Like a physical horizon, it has a scope of reference. As a function of its parameters of constraint, it determines a range of objects for which it provides a basis of identification.
- ◆ It has the property of *unboundedness*, in a fashion analogous to a physical horizon. That is, its range of inclusion may in principle be extendable (a range we have called a ‘Leibniz boundary’).
- ◆ It establishes a *modal limit* of possible objects to which it enables reference.
- ◆ It often has a *delusion-inducing character*, prompting those who employ it to trespass beyond its modal limits.

- ◆ It possesses a *metalogical horizon* which is such that, when attempts are made to overstep that horizon, projection results. A metalogical horizon determines not only limitative boundary conditions of possible knowledge, but also the limitative boundary conditions of the very nature of possibility, of meaning and reference, of identifiability, of intelligibility, and of ontology. A metalogical horizon has boundaries that become detectable “symptomatically,” i.e., those boundaries become reactively evident as a result of the occurrence of projections.

In those cases in which philosophers and logicians have variously expressed the notion of “conceptual unreachability,” we shall assess their intended use of that notion by evaluating it as a function of the above referential limitations. In the sections of this chapter that follow, we will consider a variety of ways in which the notion of “conceptual unreachability” is involved in frequent and persistent claims that have been made about the limitations of truth, knowledge, thought, and the questions we may meaningfully ask.

## **20.2 The finitude of what anyone knows and the finitude of the totality of human knowledge**

It is universally accepted (though possibly not by Faust) that the extent of any person’s knowledge is limited, and in this sense the extent of what any given human being knows is finite. When a claim like this is made, the implicit reference frame that is commonly presupposed serves as a broader context that extends beyond the limits of the individual’s range of knowledge; in other words, a more comprehensive reference frame is presumed relative to which the extent of the individual’s knowledge can be recognized as proportionally restricted and finite. Alternatively, the claim that what anyone knows is finite frequently supposes that the extent of that individual’s knowledge can potentially be increased. These are two evident ways in which an individual’s range of knowledge may meaningfully be described.<sup>236</sup>

---

<sup>236</sup> In {6.6}, I examined J. L. Mackie’s concept of operational self-refutation. In connection with the claim that “There are truths which I do not know,” it is of interest to include his view:

The modest statement ‘There are truths which I do not know’ is not self-refuting in any way. But as soon as I try to specify explicitly one of these truths, describing it as the truth that  $x$ , I am in difficulty. If I say ‘That- $x$  is a truth which I do not know’, then by calling this a truth I have committed myself to saying also that I believe that  $x$  and have reason to do so, and hence that I know that  $x$ . Thus ‘That- $x$  is a truth which I do not know’ is operationally self-refuting. This difficulty arises with any attempt to

However, when the notion of “the finitude of the totality of human knowledge” is introduced, we need to ask how the intended “finitude” and “totality” are to be understood. By definition, when “the totality of human knowledge” is considered, it is not possible, as it was in the case of the individual finite knower, to presume an implicit, more comprehensive reference frame that includes knowledge that extends beyond the limits of the totality of human knowledge. Were such knowledge to be presumed, it would evidently fall within the intended totality. If in this context a proponent of “The Far Side” were to insist that he or she intends to refer to putative knowledge that “lies beyond” the boundaries of the totality of human knowledge, a self-undermining projection results: The proponent is forced to make use of the referential capacity of his or her own reference frame in a fashion that trespasses beyond its own metalogical horizon. Such an attempt makes use of a presumed framework of reference in order to refer beyond that framework in a manner that precludes that the referential resources of that framework are any longer available. To avoid this projection, what we commonly understand by “the finitude of the totality of human knowledge” must instead be recognized to consist purely in the anticipated prospect that the sum total of finite human knowledge at any given time can potentially be extended in finite increments, so that relative to any moment of time, that sum total may, it is anticipated, be increased, but it always remains finite.

The finitude of knowledge is often expressed by:  $(x)(\exists t)\sim Kxt$ , read as: “for everyone, there are unknown truths.” Unless this is understood in the sense just proposed, it is clearly projective. Additionally, if the order of the quantifiers is reversed,  $(\exists t)(x)\sim Kxt$  is inherently projective: “there is a truth not known by anyone.” A metalogical precondition of the meaningfulness of this proposition is that a reference frame be available in terms of which the concept of truth can be employed and in terms of which a putative unknown truth can be identified. But in the case in which it is claimed that there is a truth such that for everyone it is unknown, the basis for such a reference frame is denied; no one, then in principle, has such a reference frame available, and so the *identifiability* of such a truth becomes impossible by virtue of the conditions set by the projective proposition itself.

Here, it may be useful to recall what we have observed in earlier chapters, that the identifiability of an object of reference may be specific, or it may be vague or rule-determined. Reference to an incompletely defined object that is not specifically identified still qualifies, in the terms of this study, as a form of identification: It comprises a form of general nonspecific identifying reference

---

set limits to our knowledge in this way, by actually presenting items that lie beyond the limits. (Mackie, 1964, pp. 199-200)

which describes an object of reference only in terms of the satisfaction of characteristics which perhaps could, under certain circumstances, serve to establish the specific identity of that object. In this sense, it is not as though the meaningfulness of the proposition “there is a truth not known by anyone” requires that such a truth be produced, but rather that its identifiability in principle be possible, which the proposition’s putative meaning projectively undermines. The very basis for such identifiability is withheld. We cannot, in principle, speak meaningfully about “truth” when a framework necessary for the possible meaningfulness of “truth” is ruled out.

We should more clearly express the finitude of the individual knower by:  $(x)Kxp \supset p \in \{S\}$ , where  $\{S\}$  is a finite set of true propositions, and read as: “if  $p$  is known by anyone,  $p$  is one of a finite set of true propositions.” In this way, built into the antecedent of the conditional is, so to speak, the implicit presumption of a reference frame in terms of which  $p$  is known.

### 20.3 The incompleteness of knowledge

Closely associated with the claim that what anyone knows is finite is the claim that his or her knowledge is *incomplete*—in other words, that, for any given knower, there is some true proposition that he or she does not know. This situation can be understood in terms similar to those we used in the previous section to describe ways in which an individual’s range of knowledge may meaningfully be expressed. The knowledge of an individual knower is evidently and meaningfully incomplete if there is a proposition that is known to be true by others whose scope of knowledge includes true propositions not known to the individual in view. Alternatively, an individual’s knowledge may be said to be incomplete if the range of that person’s knowledge can be extended.

However, there is a propensity among some philosophers and logicians to claim that the sum total of human knowledge is incomplete, and here, once again, we encounter a projectively self-undermining claim. No matter what item of knowledge is added to the sum total, those who make this claim may doggedly persist in maintaining that it is not the finite limit (which I have called a ‘Leibniz boundary’) that he or she insists upon “going beyond,” but rather the metalogical horizon of the sum total of human knowledge. This putative attempt to reach “The Far Side” cannot, in principle, be meaningful for the reasons given in the previous section.

### 20.4 The unlimitedness of our ignorance

“There exist an unlimited number of truths about which we are ignorant” is a claim that can be reduced to the previously considered claims of the finitude of knowledge and its incompleteness. To assert that an individual’s ignorance is unlimited, or that humanity’s ignorance is unlimited, has two evident putative purposes: to claim that the scope of an individual’s or of the human species’ knowledge can potentially be extended indefinitely, and/or to advance the set of claims that there exist an unlimited—i.e., indefinitely “discoverable” or denumerably infinite—number of truths of which the individual or the species has, or can have, no knowledge. The first claim is speculative but meaningful; the second group of claims is rife with projections which we have already discussed in this chapter, while adding to them is the projection of discovery examined in the previous chapter. A de-projective understanding of the “unlimitedness of our ignorance” leads to no more than the assertion of a prospective belief that the scope of an individual’s or of the species’ knowledge can be extended indefinitely, which is of course meaningful, whether or not it is true.

### 20.5 Thinking beyond the limits of thought

I ask you to consider an object which is not being conceived. You do so. Indeed, in the relevant sense you can conceive of anything that is referred to by a simple grammatical noun-phrase of English, just because you understand it.

$C\bar{x} \rightarrow Cx$  [a selected object is conceived that is inconceivable].

Thought can, indeed must, ... think beyond its own limits.

The thesis is that there are certain limits to thought... Thought, as it were, thinks of these limits, and, in doing so, trips over them; it cannot help but do the impossible and go beyond them (too).

– Graham Priest (1991, pp. 362, 363, 369)

The four claims above are similar in that their propounder found nothing about them that renders them devoid of possible meaning. Priest flatly asserts that, for any grammatically well-formed phrase, it is possible to conceive of anything to which it refers, but he never justifies this belief. We immediately

ask whether *any grammatical phrase* actually does refer, or even *can*. If a phrase cannot in principle refer because the conditions it sets up preclude that it can possibly refer, then we have already answered the question whether it is possible to conceive of that to which it refers: It not only *does not* refer, it *cannot*. Projective phrases embedded in the above claims not only do not refer, in principle they *cannot*, for they undermine their own possibility of referring.

In a book that sought to expand his 1991 paper, Priest (1995/2002) re-affirmed claims like this, among them:

[T]he limits of thought are boundaries which cannot be crossed, but yet which are crossed.

[In his discussion of Berkeley's thought:] It is conceived that there is something that is unconceived; or, on the modal interpretation: it is conceivable that there is something inconceivable. Both of these are unproblematically true since I, for one, conceive such things. Even to suppose them as the first step of a *reductio* requires one to conceive them (*they do not even have to be possible*).

[W]hen Sextus claims in *Outlines of Pyrrhonism* that it is impossible to assert anything about things beyond appearances, he would seem to be asserting just such a thing; and when he argues that no such assertion is justified, this must apply to his own assertion as well.

– Graham Priest (1995/2002, pp. 2-3,  
73 (*italics added*), 252)

For any reader who has come this far in the present study, brash assertions like these should elicit a response that demands a critical evaluation in terms of the referential preconditions that would need to be satisfied in order for these claims possibly to refer and to possess possible meaning.

I do not propose to point in more detail to flaws in Priest's claims, and will leave that exercise to the reader. Instead, it will be enough here to give some abbreviated idea of the context in which Priest's assertions are made. Like the present author, he has had a deeply rooted interest in conceptual and theoretical limits; unlike the present author, he has sought to make a case for contradiction-tolerant and also contradiction-asserting logics, which he has

respectively called ‘paraconsistent’ and ‘dialetheic’. Priest has been aware of some of the most fundamentally inconsistent positions taken by philosophers (e.g., Kant’s claims and arguments concerning phenomena as opposed to noumena<sup>237</sup>). He has directed attention to the pattern of inconsistency which the thought of such philosophers makes clear when they have formulated views that concern the limits of thought and knowledge, and he has posed the important question:

Now, why does this striking pattern occur again and again? ... [T]he simplest answer is that when people are *driven* to contradictions in charting the limits of thought, it is precisely because those limits are themselves contradictory. (Priest, 1995/2002, p. 252, italics added)

It is an important philosophical, and also psychological, question why people are compelled (Priest used the equally appropriate word ‘driven’) to transgress the horizons of their frameworks of reference. Priest does not give an answer why people are like this; instead, he tries to answer the question why their efforts have led to inconsistencies. These are two very different questions. His answer to the latter question is that conceptual limits, such as the limits of what can be thought, are in certain ways “contradictory,” a conclusion that supports his objective, under certain circumstances, to legitimate contradictions.

To accomplish this end, however, he was forced to make claims like those I have quoted above, claims which, though they employ grammatically well-formed phrases, are metalogically self-referentially inconsistent and therefore meaningless. It should be a straightforward matter by now for the reader to see in each instance how “the rug is pulled out from beneath each claim’s putative meaning.”

This is not a matter in which the “limits” of what can be thought are “contradictory.” It is rather a thesis that suggests that it is possible, conceivable, or meaningfully expressible somehow to reach “beyond” those “limits” by requiring us to employ reference in ways that undermine its very possibility.

We shall reserve the still-unanswered question why people are compelled or driven to engage in projections for consideration later on.

---

<sup>237</sup> Priest (2007, p. 176).



### 20.6 Expressing the inexpressible: Reaching beyond the limits of language

[Commenting on the thought of Nārgājuna:] Linguistic expression and conceptualisation can express only conventional truth; the ultimate truth is that which is inexpressible and that which transcends these limits. So it cannot be expressed or characterised. But we have just done so.

– Graham Priest (1995/2002, pp. 267-268)

[W]e could only say things about the world as a whole if we could get outside the word, if, that is to say, it ceased to be for us the whole world. Our world may be bounded for some superior being who can survey it from above, but for us, however finite it may be, it cannot have a boundary, since it has nothing outside it. Wittgenstein uses, as an analogy, the field of vision. Our field of vision does not, for us, have a visual boundary, just because there is nothing outside it, and in like manner our logical world has no logical boundary because our logic knows of nothing outside it.... In logic, therefore, we cannot say, there is this and this in the world, but not that, for to say so would apparently presuppose that we exclude certain possibilities, and this cannot be the case, since it would require that logic should go beyond the boundaries of the world as if it could contemplate these boundaries from the other side also. What we cannot think we cannot think, therefore we also cannot say what we cannot think.

– Bertrand Russell (1961/1922, p. xviii)

The putative notion that what cannot be expressed can nonetheless be expressed is, on the surface at least, self-evidently self-contradictory. The allegedly meaningful notion of what is specifically “inexpressible *in language*” is a more specialized instance of the general “inexpressibility thesis” which includes inexpressibility even in thought.

In the two quotations given above, their authors are in clear disagreement. Priest appears to be asserting that in the very act of uttering the words ‘the ultimate truth is that which is inexpressible and that which transcends these limits’ he has just sought to express, and has succeeded in expressing, in words what words cannot possibly express. To maintain this is not to maintain

the contradiction referred to in the previous paragraph, it is to assert a possibility under a set of conditions that *rule out* that possibility. To assert what Priest has asserted is therefore to assert a projection.

Russell's related thoughts on the same subject are more coherent, and yet Russell, too, is insufficiently careful from the standpoint of metalogical consistency when he claims "Our world may be bounded for some superior being who can survey it *from above*, but for us, however finite it may be, it cannot have a boundary, since it has nothing *outside* it." The terms I have italicized do, of course, make use both of our ordinary conceptual framework and of the English language, but Russell's claim projects the usage of those words beyond that conceptual framework in terms of which they have possible meaning. He similarly falls victim to projection when he states "our logical world has no logical boundary because our logic knows of nothing *outside* it," as he does when he suggests that the inexpressibility thesis would "require that logic should go *beyond* the boundaries of the world as if it could contemplate these boundaries from the *other side also*."

What both Priest and, certainly to a lesser extent, Russell failed to recognize and state in explicit terms is that it *can make no possible sense* to attempt to refer "beyond" or "outside" or on "the other side" either of what we can think or of what we can say. Inexpressibility, whether in language or in thought, is, to recall Clifton Fadiman's words, "to search for a meaning on the scrapheap of popular misuse." But it is, of course, more than this: It is a good deal more, in a logically more destructive way: It is to search for an *impossible meaning*. It is to undertake a search on behalf of a metalogically self-undermining delusion.

Passages like those I have quoted from Priest and Russell underscore the need for us to be cognizant of limits that do not have boundary lines of demarcation, boundaries for which it *can make no sense* to refer to an "outside" with respect to such boundaries. Such putative references as we find in the passages quoted at the beginning of this section can, in principle, have no meaning insofar as attempts are made to employ the word 'outside' while intending to "go beyond" the boundaries of the term's possible meaning. To attempt to do this is—and *this* claim is metalogically necessary—to employ the word 'outside', and the familiar conception of "an outside," in ways that are, in principle, meaningless. In a later section in this chapter, we shall review how it is possible, in a fashion not devoid of sense, to characterize such boundaries.

### 20.7 Fitch's Theorem

A considerable amount of attention by logicians and philosophers has been attracted by a few brief paragraphs that formed part of a short paper by Frederic Brenton Fitch (1963). The paper's objective was to provide a logical analysis of a small group of value concepts. However, a few pages into the paper, Fitch stated two very terse theorems (which he called Theorems 4 and 5) for which he provided accompanying proofs. These results began (though it took more than a decade) to take on an interest of their own. The two theorems have led to well over a hundred publications since then. I do not propose to discuss this mass of publications,<sup>238</sup> but rather to consider Fitch's two theorems and their proofs in abridged form within the framework of the metalogic of reference.

The pair of theorems and a small group of rules of derivation lead, on the one hand, in the judgment of some logicians, to a result that has come to be called 'Fitch's Theorem', which states that there are unknowable truths, e.g.,  $(\exists t)(x)(\Box \sim Kxt)$ . On the other hand, the two theorems are thought to lead, by logicians who do not accept Fitch's Theorem, to a result that has been named 'Fitch's Paradox of Unknowability' according to which if all truths are in principle knowable, then all truths are in fact known, e.g.,  $(p)(p \supset \Diamond Kp) \vdash \Box(p)(p \supset Kp)$ .

Whether there are or are not truths that, in principle, are unknowable, or whether it even is meaningful to consider such a possibility, is clearly a topic to which the present study's concept of metalogical horizon would seem to apply. Let us see if it does, and if so, how.

Fitch's Theorem 4 states "For each agent who is not omniscient, there is a true proposition which that agent cannot know" (Fitch, 1963, p. 138). For this theorem, Fitch gave a very short proof which makes use of the supposition " $p$  is true but not known by the agent." From this supposition, Fitch reached the conclusion that "there is some true proposition which cannot be known by the agent" (p. 139).

This then leads to Theorem 5, which states: "If there is some true proposition which nobody knows (or has known or will know) to be true, then there is a true proposition which nobody can know to be true" (p. 139)—in other words, an unknowable truth.

There now exist many variations of the alleged proof and many variant reconstructions and interpretations of Fitch's result, a discussion of which is

---

<sup>238</sup> A few sample informative publications that will lead the reader further include Beall (2000), Edgington (2010), Fara (2010), Fitch (1963), Mackie (1980), Rescher (2005, 2009), Routley (2010/1981), Salerno (2009), Williamson (2000).

beyond our purpose and not relevant here.<sup>239</sup> Instead, let us look more closely at Fitch's original pair of theorems. They involve these claims:

For each agent who is not omniscient, there is a true proposition which that agent cannot know. (i)

$p$  is true but not known by the agent. [supposition] (ii)

There is some true proposition which cannot be known by the agent. (iii)

If there is some true proposition which nobody knows (or has known or will know) to be true, then there is a true proposition which nobody can know to be true. (iv)

Among the concepts that evidently are used to make these claims are the concept of "an agent who is not omniscient," the concept of a proposition "not known by the agent," the concept of "a true proposition which that agent cannot know," and the concept of "a true proposition which nobody knows (or has known or will know) to be true." Each of these notions calls for referential analysis so that we may recognize the preconditions of possible meaningfulness of each. We shall use this occasion as a simple exercise in the variety of analysis with which we have been concerned in this study; the reader will note the distinctive level of analysis than would be found were we instead engaged in a purely logical analysis of Fitch's theorem.

By "an agent who is not omniscient" let us stipulate that this means, for Fitch, two things, one of which is not explicitly stated: that the agent in question possesses only finite knowledge, and that there exists *other* knowledge not possessed by that agent (that is, other true propositions not known to be true by that agent). In this presupposed context, (i) claims that there is a true proposition which that agent cannot know.

In Fitch's proof, the supposition (ii) is made that some proposition  $p$  is true but that  $p$ 's truth is not known by the agent, whom we shall call 'agent<sub>0</sub>', and let us call agent<sub>0</sub>'s own reference frame ' $f_0$ '. To be possibly meaningful, what *must* supposition (ii) presuppose? It must presuppose that there is a suitable reference frame,  $f_1$ , which allows some agent<sub>1</sub> (who is not the same as agent<sub>0</sub>), in principle, to recognize that  $p$  is true. (For otherwise how could it,

---

<sup>239</sup> For Fitch's own formulation of his result, see Fitch (1963), and for alternative representations of Fitch's result, commentary, and criticism, see the examples of sources listed in the previous note.

in principle, be meaningful to make the supposition that  $p$  is true? A reference frame such as  $f_1$  is necessary in order to meet the condition that  $p$  is true but is not known to be true by agent<sub>0</sub>. That supposition must make sense if it is to be used in the course of Fitch's proof, and to make sense, it presupposes a possible appropriate reference frame in terms of which the concept of "truth" can be employed so that  $p$  can be claimed to be true.)

What follows from this in Fitch's proof is (iii), that there is some true proposition which *cannot* be known by agent<sub>0</sub>. We notice that this is a different kind of claim, relating not to what agent<sub>0</sub> doesn't know, but to what it is not *possible* for agent<sub>0</sub> to know. In this latter case, agent<sub>0</sub> is *unable* to know that  $p$  is true, although it is. To be possibly meaningful what *must* this latter claim presuppose?

The framework necessary to know that  $p$  is true is not necessarily a framework that has the referential resources to know that  $p$  cannot be known to be true by agent<sub>0</sub> by means of framework<sub>0</sub>. Let us call the latter frame of reference ' $f_2$ '; we notice that it is a modal framework, i.e., enabling reference to what it is not *possible* for agent<sub>0</sub> to know by the means afforded by that agent's framework<sub>0</sub>. We therefore do not assume that  $f_1$  and  $f_2$  are the same framework, although such a dual-purpose framework is possible. We then have the situation in which (iii) presupposes a different reference frame,  $f_2$ , so that when  $f_2$  is employed by a different agent<sub>2</sub>, that agent can know that there is a true proposition that *cannot* be known by agent<sub>0</sub>. Agent<sub>1</sub> plays the role of knowing in  $f_1$  that  $p$  is true; agent<sub>2</sub> plays the role of knowing in  $f_2$  that  $p$  cannot be known by agent<sub>0</sub> using  $f_0$ . Both reference frames,  $f_1$  and  $f_2$ , employed by agent<sub>1</sub> and agent<sub>2</sub>, are required for supposition (ii) and assertion (iii) to be expressed with their intended senses.

Now, (iv) states: "If there is some true proposition which nobody knows (or has known or will know) to be true, then there is a true proposition which nobody can know to be true." What conditions must be granted in order for this claim possibly to be meaningful? The antecedent of the conditional assumes hypothetically that there is a proposition not known by anyone to be true, let us call it ' $P$ ', which nonetheless is true, but since nobody knows  $P$  is true (or has known or will know that  $P$  is true), there is no reference frame relative to which  $P$  can be identified and known to be true. Assertion (iv) is, then, metalogically self-undermining. It sets up conditions that cannot, in principle, be satisfied.

Of the many publications that have been devoted to an analysis of Fitch's Theorem, none, to my knowledge, has examined the referential preconditions that must be met in order for his Theorem to be stated meaningfully and proved. His reasoning can and clearly has been formalized, but such formal-

izations have not rendered explicit the referential preconditions that must be met in order for the concepts so formalized in principle to be meaningful. It should come as no surprise that a formalization, and even a formal proof that is logically compelling, is possible even though one or more of the propositions formalized are metalogically self-undermining.

### 20.8 What *can* ‘non-omniscience’ and ‘knowability’ mean?

Fitch’s Theorem has been thought by some logicians to lead to the result that there are unknown truths. In making the case to support this contention, the notion of the finitude of a knower’s knowledge is employed, often called ‘non-omniscience’ and formally expressed, e.g., by:  $(x)(\exists t)\sim Kxt$ ,<sup>240</sup> to be read as “for everyone, there is a truth that is not known.” This proposition is projective, for the reasons relating to Fitch’s claim (iv) above. Instead, the finitude of human knowledge is more appropriately expressed by:  $(x)Kxp \supset p \in \{S\}$ , where  $S$  is a finite set of true propositions, i.e., “for anyone who knows an arbitrary proposition  $p$ , it follows that  $p$  is a member of a finite set of true propositions.”

In much the same way, many logicians employ what has come to be known as the *knowability principle*, that  $p \supset \Diamond Kp$ , i.e., if  $p$  is true, it is possible to know that  $p$  is true. The supposition is then often made that, as a collective group, human beings are non-omniscient, leading to the result that there is an unknown truth (e.g.,  $(\exists p)(p \wedge \sim Kp)$ ). However, by virtue of the asserted fact that we are finite knowers, we possess no reference frame in terms of which it is possible to claim that there is a  $p$  that is true that is not known, for clearly then we cannot know this. To be able to claim that there is a  $p$  that is true, we must possess a reference frame in terms of which  $p$  can in principle be known to be true, and the availability of such a reference frame is denied by the claim that there is an unknown truth. Such a claim is, in short, projectively self-undermining.

### 20.9 Fitch’s Theorem: The de-projective result

The above short analysis leads to the following conclusion: The proposition that there are in principle unknown truths is projective. And yet, what is its putative meaning? As we’ve seen, it is plausible that the intent in claiming that there are unknown truths is to recognize that, since knowledge is finite, the scope of what is known may, with the passage of time, be enlarged. Here

---

<sup>240</sup> As in Rescher (2005, p. 68).

we find an application of the concept of Leibniz boundaries developed in {14}. It is in this sense that the scope of what we know may generally be enlarged with the passage of time, but always that scope remains finite and therefore limited.

Interestingly, Fitch, in his paper which has led to so much discussion of unknowability, explicitly recognized the potential role of time in his analysis, and yet he chose to ignore it for simplicity. As he expressed this decision:

For purposes of simplification, the element of time will be ignored in dealing with these various concepts. *A more detailed treatment would require that time be taken seriously.* One method would be to treat these concepts as a three-termed relation between an agent, a proposition, and a time. Another method would be to avoid specifying times explicitly, but rather to use a temporal ordering relation between states of affairs. This latter method might be more in keeping with the theory of relativity, in either its special or general form. (Fitch, 1963, p. 136, italics added)

Unknowability, from the standpoint of the referential analysis summarized here, is a concept that leaves open the possibility that a reference frame may, with the passage of time, come to include the knowability of truths that it does not include at present, or did not include in the past. What our analysis denies is that it can, in principle, make sense projectively to claim that there are truths, relative to a given reference frame, which that reference frame does not include. All of the truths that a given reference frame includes, at a particular time, may not be all the truths which that frame may subsequently, with the passage of time, come to include.

We shall find that other results, which complement this conclusion, will be reached in connection with a variety of other epistemologically important topics, e.g., relating to the problem of the external world, relativity physics, and quantum theory.

### **20.10 Unanswerable questions: Erotetic intractability**

[W]e can ... contemplate the prospect of globally intractable questions such that nobody (among finite intelligences, at least) can possibly be in a position to answer them.... These questions have an appropriate answer, but for reasons of

general principle no one—no finite intelligence, at least—can possibly be in a position to provide it.

– Nicholas Rescher (2005, p. 92)

By ‘*erotetic intractability*’ is here meant the impossibility, in principle, that certain putatively meaningful questions can be answered. Questions may be unanswerable for a variety of reasons, among them: lack of acquaintance or knowledge of relevant facts, lack of physical or informative resources, unavailability at the time of solutions that respond to a question’s request, as well as many other reasons, including a simple failure to understand the language in which a question is expressed, the technical terms it employs, previous knowledge it presumes, etc. However, *in-principle unanswerable questions* are not unanswerable for any of these or related reasons.

Many philosophers have asked unanswerable questions, and some were probably unaware that no answers to them were, in principle, possible. But some philosophers have claimed that certain questions can be both impossible to answer and yet be meaningful. Is this the case?

Rescher has been an advocate of the meaningfulness of in-principle unanswerable questions. Two of his books (Rescher, 2005 and 2009) claim such questions are meaningful and provide examples of questions of this sort. Here are a few representative passages:

[C]onsider such questions as: “What is an example of a problem that will never be considered by any human being?” “What is an example of an idea that will never occur to any human being?”... (Rescher, 2005, p. 91)

There are sound reasons of general principle (the potential infinitude of problems and ideas; the inherent finitude of human intelligence) to hold that the items at issue in these questions (problems that will never be considered; ideas that will never occur) *do actually exist*.... (Rescher, 2005, p. 92, italics added, and repeated verbatim in Rescher, 2009, p. 9)

An example of ... globally unanswerable questions can be provided by nontrivial but inherently uninstantiable predicates along the lines of “What idea is there that has never occurred to anybody?”...

There undoubtedly *are* such items, but, of course, they cannot be instantiated, so questions that ask for examples here are inherently unanswerable....



With answer-possessing but unanswerable questions it accordingly must transpire that the answer that, abstractly speaking, *has to be there* is one that cannot possibly be specified by way of particularized identification. (Rescher, 2005, pp. 92- 93, italics added)

A few years later in his book on the same subject, Rescher continued to maintain these bluntly asserted beliefs: “[T]he questions that will preoccupy [me] are not just questions we cannot answer, but questions which, in the very nature of things, no one can possibly answer (Rescher, 2009, p. 2).

What—we are compelled to ask—is the nature of a meaningful question of this particular kind that requests a true and informative answer?

There are, to be sure, numerous varieties of questions, a catalogue of which is not our present interest. What interests us here is of course that variety of question which is involved when philosophers pose questions like those by Rescher above. It is—putatively—a species of question that requests an answer that can be known to be true. This variety of question is much like a problem as defined in the previous chapter: A problem directs attention to a gap between a set of initial conditions and a desired goal. But the alleged in-principle unanswerable questions are considerably different; they may appear to set a problem to be solved, but they do not. Let us see how this is the case.

Charles Sanders Peirce observed long ago that a question is a “rational contrivance or device” that expresses a request for an unknown that is sought after and indicates “what has to be thought about something in order to satisfy” that request (Peirce, 1931-1935, vol. 3, p. 414). R. B. Angell sharpened this observation: A question focuses attention

... on something that is missing; it points, as it were, to something not completely specified, urging one on to complete the specification.... [E]very question ... *if significant at all*, also gives us some knowledge or some partial specification of what it is we are looking for. From the question itself we both (a) know some characteristics which will have to be found in the answer, and (b) some kinds of characteristics which remain to be specified and found. (Angell, 1964, p. 457, italics added).

Combining these observations by Peirce and Angell,<sup>241</sup> we might formulate three necessary conditions which a meaningful question of the indicated kind

---

<sup>241</sup> I am grateful to Knight (1967) for bringing these authors’ papers to my attention.

must satisfy: Such a meaningful question must (i) express a request for an unknown, (ii) provide information that communicates in some sense what conditions a satisfactory response would have to meet, and (iii) indicate, again in some sense, what would have to be done in order to obtain such a satisfactory response. A question that does not meet these three conditions would not, if one accepts conditions (i)-(iii), be considered meaningful.

These conditions, which are likely to be accepted by the majority inquirers, would seem to stand on their own as self-evidently convincing. But it would be good to have stronger, more compelling assurance of their persuasiveness.

Let us agree to call the particular variety of truth-inquiring question on which we shall focus a 'T-question'. On its most basic information-theoretic level, a meaningful T-question formulates a request for a true and informative reply. The information such a meaningful question asks for may be specific or non-specific; a clearly identified answer or only a vaguely identified answer may be sought. A meaningful T-question may sometimes need to be clarified and refined, and the conditions it sets for possible answers to it may need to be made explicit or explained. But, whether the information sought is specific or not, in either case, a meaningful question of the variety that concerns us presupposes one, and sometimes more than one, frame of reference in terms of which such information can be searched for and supplied. Such a frame of reference is often implicitly implied by the question and the context in which the question is posed, and in this sense, the communication of a meaningful T-question requires an appropriately informed respondent: He or she may not know the answer to the question, but he or she needs to be able to understand what, to some specific or non-specific degree, is being asked. To do this, he or she must recognize, to some degree, the frame of reference presupposed by the question. In this sense, a meaningful T-question presumes an appropriately "equipped" recipient. The question must be *intelligible* in this way; its possible meaning as a question requires this.

Erotetic intractability *amputates* a question from any possible reference frame in terms of which it *can be* answered. The putatively meaningful content expressed by the words used to formulate an in-principle unanswerable question is "severed" from the system of reference that *appears* to be established by those words. What we see is a kind of conceptual "sleight-of-hand" in which words are used to undercut the very preconditions that must be satisfied in order for them to refer as apparently intended. The words, so to speak, appear to point to something, but, at the same time, what those words are taken to mean undermines the possibility that they *could* point at all in the manner intended.

Erotetic intractability inevitably brings with it metalogically projective delusional reference. Questions like the ones quoted above from Rescher are therefore projective and can have no possible meaning. They appear, initially and before reflective analysis, to possess a putative meaning, but although words and phrases are employed that have meaning individually, taken together those words and phrases express a delusion-inducing intention that is metalogically self-undermining. Unlike a general problem that points to a gap between an initial state and a desired goal, the question's "initial state"—the set of conditions expressed by an in-principle unanswerable question—establishes as a putative "goal"—a response which putatively might be true and informative—but which undermines its own possibility of reference and has no possible meaning.

Erotetic intractability plagues a number of important philosophical questions. We shall occasionally have the need to renew our acquaintance with in-principle unanswerable questions in subsequent chapters.

### **20.11 The philosophical and psychological compulsion to trespass beyond horizon boundaries**

In this chapter, we have seen how philosophers and logicians have tried to reach for the conceptually unreachable. We have discussed this phenomenon in connection with

- ◆ the finitude of knowledge
- ◆ the incompleteness of knowledge
- ◆ the unlimitedness of our ignorance
- ◆ thinking beyond the limits of thought
- ◆ expressing the inexpressible
- ◆ Fitch's Theorem
- ◆ unanswerable questions

In each of these cases of projection, it does not seem to be an exaggeration to say that a "compulsion" (as noted by Kant)<sup>242</sup> or a "drive" (to use Priest's term) is at work in the minds of those who insist on reaching beyond metalogical boundaries. Philosophers and logicians who express this insistence are legion. They surely respond to what they perceive as a conceptual need to make projective claims, and in doing this they do, we may suspect, also respond to a psychological need that manifests itself in this reaching for what is

---

<sup>242</sup> We recall that the phrase he used was "*fühlt sich nicht notgedrungen*" (see {14}).

in principle unreachable. Rescher is honest when he admits to “a *conviction* that there is more to reality than we humans do or can know about” (Rescher, 2005, p. 101, italics added). This *conviction*, like so many others, is strong and highly resistant to criticism, as early chapters in this book have discussed. If, in fact, projective thinking is essentially delusional in nature—if projections lead to actual *delusional thinking*<sup>243</sup>—this goes a long way in explaining the recalcitrance of philosophers through the ages to curtail their drive to *overstep* the horizons of their frameworks of reference.

So often, as we have seen in this chapter alone, the urge to attempt to reach for the unreachable takes the form of attempting “to say the unsayable,” “to think the unthinkable,” “to refer to that which cannot be referred to,” etc. Perhaps most frequently, the conceptual *drive to over-reach* derives from the epistemologically naive human belief that “reality must be more than we know,” and the class of beliefs related to this. Again to quote Rescher:

[T]here is a reality out there that lies beyond the inadequate gropings of mind.... Every sector of reality has features lying beyond our present cognitive reach—at any “present” whatsoever.... [W]e cannot justifiably equate reality with what can, in principle, be known by us and thereby, in principle, be expressed in our language.... [W]e cannot possibly articulate, and thus come to know explicitly, “the whole story” about things. The domain of fact inevitably transcends the limits of our capacity to express it, and a fortiori those of our capacity to canvass it in overt detail. (Rescher, 2005, pp. 102, 113, 121)

And, yet, inspirational and buoyant philosophical claims like this aside, no philosopher has ever been able to demonstrate that the above propositions, to which so many are *convinced*, are actually true, or even knowable—and, in the framework of the present study, we should add: even *possibly* meaningful. Such horizon-transgressing claims are characteristically made categorically, with determination, often unhesitatingly and cavalierly, as if the degree to which they are adamantly asserted has, in itself, persuasive force. Here, the projective conceptual drive and the psychological compulsion intertwine and mutually reinforce one another.

---

<sup>243</sup> Projections also can lead to *delusional behavior*, as observed in Bartlett (2005, Part III; and 2011, Chap. 9 and *passim*.)

### **20.12 The need to resort to suggestive—e.g., rule-based or property-based—reference**

From a strict epistemological standpoint, there is nothing inherently objectionable about the simple expectation that if one were to continue sailing toward the visible physical horizon, the range of what one sees would very likely expand. Similarly, there is nothing epistemologically objectionable about the simple expectation that one should be able to continue a series of integers. Such expectations that Leibniz boundaries may be extended do not comprise the central problem this study seeks to examine and solve. However, we become so habituated to our beliefs in the potential extendability of Leibniz boundaries that we come to rely on those beliefs in our practical lives. We develop a predisposition to expect that “more” will be revealed as a Leibniz boundary is extended, but then, unfortunately, we take a rash step further and transgress beyond the possible range of our frameworks of reference, and then we fall victims to asserting projectively self-undermining and meaningless propositions.

It is difficult to curtail the expression of this projective human predisposition, despite attempts like those made in this book. What we appear to need, conceptually and psychologically, is a set of meaningful ways in which we can identify objects of reference in a fashion that is intelligible and coherent, in a manner so that one knows what one is talking about, while still possessing the means to refer to what we believe are incompletely defined, incompletely represented, or incompletely experienced objects.

But we do already have, ready-to-hand, such means of identifying what we consider to be incompletely given objects of reference. We make recourse to general nonspecific forms of reference that identify an object of reference only in terms of the satisfaction of characteristics which would, under certain specific circumstances, serve to establish the identity of that object more clearly or precisely. I shall call these nonspecific forms of reference ‘rule-based’ or ‘property-based’ reference. They enable us to refer to objects in as yet nonspecific ways, according to a set of usually implicit rules or properties which, when they are relied upon, provide a basis for defining, representing, or experiencing objects that we believe are incomplete in any of the above-mentioned senses.

For example, we may wish to view a three-dimensional physical object from another perspective, and so we walk around the object, habitually expecting that as we do this, our Leibniz-bounded field of vision will change and new visual information about the object will be available to us. Implicit in this pattern of anticipation and behavior is, one could say, a set of rules for dealing with three-dimensional objects. (Such a set of rules is made explicit,

for example, in the code written so that an artificial intelligence can similarly interact with physical three-dimensional objects.) In much the same way, we may wish to know more about the works of a particular well-known author, say author Y; we go the library with the anticipation of finding other works by Y; we borrow some of these; open several volumes; and can now learn more about that Y's thought. Again, we have extended a different sort of Leibniz boundary. Implicit in this pattern of anticipation and behavior, again one could say, is our reliance upon a set of defining properties that direct our attention and activity to the end of "learning more about author Y's work."

Such examples make evident that *it is only relative to the reference frames we employ that the extendability, the identifiability, the meaning, and even what we understand to be the reality of that to which we refer, are possible* (see {18}). To wish projectively to claim that "there is more than this"—or to deny this and assert the negation of this projection—is to fall victim to projective delusion. We shall find this exemplified again in the next chapter when we examine the subjects of realism, idealism, and the problem of the external world.

### 20.13 Unbounded yet topologically enclosed frameworks

[G]reat stress is laid on the limitations of thought, reason, and so on, and it is asserted [by some] that the limitation *cannot* be transcended. To make such an assertion is to be unaware that the very fact that something is determined as a limitation implies that the limitation is already transcended. For a determinateness, a limit, is determined as a limitation only in opposition to its other in general, that is, in opposition to that which is *free from the limitation*; the other of a limitation is precisely the *being beyond it*.

– G. W. F. Hegel (1969/1812, p. 134)<sup>244</sup>

Thought can, indeed must, therefore, think *beyond its own limits*.

– Graham Priest (1991, p. 369, italics added)

---

<sup>244</sup> To avoid bestowing on this often-quoted passage a value it does not merit, we should be reminded that here Hegel was writing about what he called '*the ought*'—that is, "the moral obligation to transcend limitations" (Hegel, 1969/1812, p. 134). The passage is often quoted out of context: It is a fantasy of hopeful attribution to think it expresses an epistemological or metological insight.

The objective of this chapter in conjunction with the chapters before it has been to show that assertions like those made more than 200 years ago by Hegel and more recently by Priest are not only fundamentally mistaken, but they are expressions of a form of conceptual delusion that is provably meaningless. A step closer to the conclusions that we have reached was made in a short passage written nearly a century ago by Wittgenstein, when he wrote:

[T]he aim of the book is to draw a limit to thought, or rather—not to thought, but to the expression of thought: for in order to be able to draw a limit to thought, we should have to find both sides of the limit thinkable (i.e. we should have to be able to think what cannot be thought).

It will therefore be only in language that the limit can be drawn and *what lies on the other side of the limit will simply be nonsense*. (Wittgenstein, 1961/1921, p. 3, italics added)

Unfortunately, Wittgenstein, too, fell victim to Hegel's and Priest's projective beliefs that "in order to be able to draw a limit to thought, we should have to find both sides of the limit thinkable (i.e. we should have to be able to think what cannot be thought)." However, on the positive side, he made an advance through his claim that "what lies on the other side of the limit will simply be nonsense." But in asserting this, he was still not able entirely to shed projective thinking, for there is no possible meaning that can be associated with "what lies on the other side of the limit."

What we therefore stand in need of is a way to understand the special variety of limitation for which it is meaningless to suggest that it has, or does not have, "an outside." The concept of metalogical horizon provides a way of understanding, in meaningful, non-self-undermining terms, the variety of limitation to which thought and its expression are subject.

In {4.10}, I introduced the concept of the *self-enclosure* of reference frames. The concept of self-enclosure is central to our understanding of this special variety of limitation for which it makes no sense to claim that either it has, or fails to have, "an outside." It will be useful to summarize here what we have come to understand about the nature of self-enclosure. The following fundamental observations relating to the concept of self-enclosure have been made:

- ◆ Philosophical systems themselves possess the property of self-enclosure in the sense of a set of propositions that is "closed upon it-

self,” much like the self-enclosure of closed ideological systems of belief.

- ◆ From a phenomenological point of view, experience forms a self-enclosed “field”: However experience may be extended, it is never possible to “go beyond” its boundaries.
- ◆ In a similar way, the systems of belief of delusional psychiatric patients may possess the property of self-enclosure, exhibiting stubborn resistance to evidence that opposes those beliefs.<sup>245</sup>
- ◆ The reflexive, vertical, non-ordinal frame of reference established on the level of transcendental, maximum theoretical generality by the metalogic of reference forms a self-enclosed system.

In {10.4} the concept of self-enclosure was further elaborated. We found that:

- ◆ The mathematical concept of field may be understood as a minimalist formal model characterized by self-enclosure.
- ◆ In a parallel manner, the quantum-theoretical concept of field offers an abstract model in terms of which the elements of a system may be understood as instantiations or realizations of the field or as “epiphenomena” of that system.
- ◆ Complementing the mathematical and quantum-theoretical concepts of field, framework relativity may be characterized in terms of the self-enclosure of what we have called the ‘referential field’ that a given frame of reference establishes, and as a function of which a range of possible objects is understood to be self-enclosed. In this sense, a framework permitting identifying reference to a range of possible objects is said to be ‘*relativistically self-enclosed*’.<sup>246</sup>
- ◆ We saw that the property of the self-enclosure of referential fields has ontological consequences: Both the assertion as well as the denial of ontological framework-independence was shown to be metalogically

---

<sup>245</sup> As Appendix II shows in some detail, this stubborn resistance to evidence that opposes beliefs that individuals prefer to embrace is common also among philosophers.

<sup>246</sup> For more on the concept of relativistic self-enclosure, see Bartlett (1970, Sections 1.6, 2.1, and 2.2).



self-undermining. *To be*—in any manner that is meaningful and therefore intelligible—*is to be an object of reference, and to be an object of reference is to be an object of reference from the standpoint of an appropriate reference frame.* To claim that “to be” is more than this, or to deny this, is projectively self-undermining.

- ◆ A given framework of reference establishes a referential field in terms of which possible objects of reference can be identified. The identity of an object of reference then may be understood as an “instantiation” or “realization” of the referential field of its presupposed reference frame. Although such a field always has finitary limits, it may have Leibniz boundaries that potentially can be extended. And yet such a referential field—which has no delimited boundary lines of demarcation and is, in the sense which this study has developed, *unbounded*—is characterized by “reactive” boundaries that are made evident when attempts are made to transgress its metalogical horizon.

In {15.4}, this result was added:

- ◆ The formal structure of the method of de-projection is reflexive, possessing the property that we have called ‘recurvature’ (cf. {3, 4, 9, 14}): Its structure is self-enclosed in the sense in which transcendental argumentation exhibits a special form of circularity, and in the topological sense of possessing no “outside.”

These results, when taken together, provide a way to understand, in a coherent and non-projective way, the theoretically interesting and often perplexing nature of that special variety of limitation that has occupied us in this chapter. It is a form of limitation for which it is conceptually impossible in any possibly meaningful way to claim—as have philosophers and logicians from Kant and Hegel, to Wittgenstein, and continuing today—that there are truths that we cannot possibly know; that knowledge is necessarily incomplete; that our ignorance is unlimited; that we can think beyond the limits of thought and express the inexpressible; that, à la Fitch, there are unknowable truths, or that if all truths are in principle knowable, then all truths are in fact known; and that there are in-principle unanswerable questions. We have seen that all of these claims are afflicted by projectively self-undermining concepts.

The purpose in recognizing this metalogical fact is to make it possible, gradually and incrementally, to develop an integrated understanding that affirms framework relativity and accepts that it is only by conforming to its

metalogical requirements that we are able to develop a meaningful and strongly compelling worldview, an objective that was once, and still can be, a fundamental purpose of a rigorous, scientific philosophy.

### 20.14 Conclusion

The preceding chapters have been concerned with the preconditions that must be satisfied in order for reference and meaning to be possible. That concern, we might say, defines “*This Side*” (again to resort to metaphor). “*This Side*” comprises the whole, not merely of human experience, but the set of all possible frames of reference. When the preconditions of reference in which we have been interested are made explicit, they define the metalogical horizon of what, from the standpoint of a given reference frame, is possible and meaningful. Philosophers who have made characteristic claims about “*The Far Side*” necessarily make use of the referential resources of “*This Side*” in order to make claims and ask questions that—putatively—refer in various ways beyond the referential capacities of “*This Side*.”

In this chapter, we’ve examined a variety of ways in which attempts have been made to reach what I’ve metaphorically called ‘*The Far Side*’. We have wished to answer a central general question, which we may now express by posing it in simple terms: Is there a “*Far Side*,” or only “*This Side*”? Since the referential resources of “*This Side*” are employed seemingly to point to, suggest, or intimate a putative “*Far Side*,” what we have found is that “*The Far Side*” is wholly absorbed by “*This Side*” to such an extent that “*The Far Side*” is divested of all possible putatively meaningful existence “beyond” “*This Side*,” and so undermines itself and collapses into meaninglessness. We have found find that “*This Side*” has what we might call ‘a referential topology’ that is self-enclosed in the sense summarized in the preceding section.

In that light, it *can* make no sense—even metaphorically—to use the twin expressions ‘*This Side*’ as opposed to ‘*The Far Side*’: In creating these expressions and using them *as though* meaningful, we have engaged in conceptual sleight-of-hand, which should by now be transparent: We have, so to speak, played upon the keys of a habitually familiar ghost piano to create a paracusia, an auditory hallucination that we seem to hear as a meaningful melody.

To the extent that we are rational beings, we need instead to relinquish the compulsion to transgress the horizons of our frameworks of reference and recognize that what we now dismiss as “*This Side*” comprises “reality,” as this concept is understandable in the ontological terms described earlier in {18}.

## 21

### **The Projections of the External World, Things-in-themselves, Other Minds, Realism, and Idealism**

[I]n the sphere of abstract thought, error may reign for a thousand years, impose its yoke upon whole nations, extend to the noblest impulses of humanity, and, by the help of its slaves and its dupes, may chain and fetter those whom it cannot deceive....

[W]e ought to follow truth even although no utility can be seen in it, because it may have indirect utility which may appear when it is least expected; and I would add to that we ought to be just as anxious to discover and to root out all error even when no harm is anticipated from it, because its mischief may be very indirect, and may suddenly appear when we do not expect it, for all error has poison at its heart.

– Arthur Schopenhauer (1909/1818-1819,  
vol. 1, pp. 45-46)

**B**y far the greatest obstacle to progress in philosophical thought is the resistance by those whose unshakable adherence to common sense stands firmly in the way of a critical examination of their own beliefs. This resistance is expressed in the combative hostility aroused when their intuitively obvious common sense is questioned. Opposition to the counterintuitive has obstructed, repeatedly set back, and frequently defeated human development throughout its history. There is no disciplinary field of study that has not suffered from and been hampered by the stubborn, recalcitrant, defensive, and antagonistic reactions by both the specialist and the common person to what they feel challenges the intuitively self-evident.

By '*intuition*' I mean the set of human and non-human animal abilities to recognize, interpret, and respond to events, situations, problems, and the thoughts and behavior of others in ways which do not explicitly involve, or which efficiently circumvent, the processes of reflective reasoning. Intuition

in this sense can be a valuable survival asset in urgent situations that do not provide the leisure for deliberative or contemplative thought. But, despite this advantage, intuition imposes a very considerable liability as a result of its ability effectively to crowd out or to supersede rational responses. From the standpoint of philosophical reflection, intuition consists of a set of abilities that are primitive in the sense pointed out by Russell:

[I]ntuition ... seems on the whole to diminish as civilization increases. Speaking broadly, it is greater in children than in adults, in the uneducated than in the educated. Probably in dogs it exceeds anything to be found in human beings. But those who find in these facts a recommendation of intuition ought to return to running wild in the woods, dyeing themselves with woad and living on hips and haws. (Russell, 1972/1914, p. 34)

Often, an intuitive response is believed by the intuitive respondent to comprise a form of direct and immediate insight, one characterized by an unmistakably high degree of lucidity, convincingness, and unquestionability. The perceived clarity and utter persuasiveness of intuition have a soporific and comforting effect, contributing to an organism's felt sense of security in its world of experience, and saving the organism from the need to engage in the work of reflective thought. Since intuitive abilities have developed so as to enable them to bypass and override slower and more cumbersome rational processes, their capacity to block and undermine rational thought and its expression should not be underestimated.

Perhaps more than any other topics considered by philosophers, those of "the reality of the external world," "the existence of things-in-themselves," and "the existence of other minds" elicit among the majority of philosophers and scientists, along with the larger part of humanity, such strong, intuitively compelling responses as to make these topics stubbornly resistant to patient, rigorous, reflectively critical analysis. In subsequent chapters, we shall repeatedly see how resistance to the counterintuitive has shackled and impeded philosophical progress.

### **21.1 The projection of the external world**

We naturally believe, for example, that tables and chairs, trees and mountains, are still there when we turn our backs on them. I do not wish for a moment to maintain that this is

certainly not the case, but I do maintain that the question whether it is the case is not to be settled offhand on any supposed ground of obviousness.... As soon as the question is seriously raised whether, because we have seen them, we have a right to suppose that they are there still, we feel that some kind of argument must be produced, and that if none is forthcoming, our belief can be no more than a pious opinion....

This common-sense belief, it should be noticed, is a piece of audacious metaphysical theorizing; objects are not continually present to sensation, and it may be doubted whether they are there when they are not seen or felt. This problem, which has been acute since the time of Berkeley, is ignored by common sense, and has therefore hitherto been ignored by physicists.

– Bertrand Russell (1972/1922, pp. 77, 107)

Since Russell wrote these lines, a century has passed. During that time, much has changed in theoretical physics, specifically as a result of an increasing willingness to renounce belief in a physical reality independent of the observer. This change has come about largely as a result of developments in quantum theory. As we shall see later in {27}, the “audacious metaphysical theorizing” that Russell mentions is no longer dominant in quantum physics, although similar advances have not yet been incorporated in either common sense or in mainstream philosophical thought. The status of “the external world” doggedly persists as a much-discussed question among many philosophers. It is a question that can profit by the application of the method of de-projection.

There are two principal ways in which the concept of “the external” is applied when we speak of “the external world.” Let us look at each of these in turn.

### 1. *The external spatial world*

We start by living, each one of us, in the solitariness of our own minds and from the data given us and our communications with other minds we construct the outside world to suit our needs. Because we are all the result of one evolutionary process, and our environment is more or less the same, the constructions we make are roughly similar. For convenience

and simplicity we accept them as identical and speak of a common world.

– W. Somerset Maugham (1950, p. 526)

Maugham (1874–1965) was a British writer of fiction and plays, not a philosopher. His simple commonsense account of how we come to the notion of a shared, external world may serve as an unrefined and commonplace starting point in understanding the concept of an external spatial world.

In {10.5}, I defined the term ‘perspective’, which I shall continue to use in the present context with some slight misgiving because, in ordinary use, the word can sometimes connote a “partial view” of an implied, existing totality. Nothing of that kind is intended here: A “perspective” was defined as any one of a group of compatible reference frames in terms of which a set of objects can be identified. And by “compatible reference frames” was understood any frames of references capable of referring to the same object. Compatible reference frames or “perspectives” can be coordinated so that objects in one can be recognized as the same as objects in another. It is only in the extended sense that any one of these reference frames, taken out of its coordinated association with others, provides a basis for reference that is “part” of their collective ensemble. In this, there is no assumption that a corresponding whole “exists,” in any sense, apart from or in separation from the individual reference frames.

To illustrate this understanding of perspective, I gave an uncomplicated example of two sets of coordinates, one in Cartesian notation and the other in polar notation, chosen because they refer to “the same point” that is established by familiar transformation rules which translate coordinates of one system into those of the other and, as a result of the application of these rules, refer to what is then identified to be one and the same point. In a similar way, reference frames generally, whether or not they are formal coordinate systems, are considered to be compatible when the identity of an object given in one frame can be translated into a correspondingly identified object in the other frame in such a way that the objects identified in both frames are considered to be the same.

We rely upon the perceived compatibility of our various individual spatial frames of reference to accomplish the foregoing, i.e., when we identify a physical object that is considered to be the same from our different individual perspectives.<sup>247</sup> In the case of a physical object, such perspectives are most

---

<sup>247</sup> Our largely linear language does not permit us to talk simultaneously and effectively about distinguishable but interrelated topics; that of “other minds” is considered in a later section of

commonly visual, they are sometimes tactile, and sometimes auditory, or perhaps olfactory or gustatory. Each of these perspectives provided by our sensory means supplies us with a basis to refer to objects in a correlational way—correlating what I see with what you see, etc. For the most part the processes that we utilize to make these correlations do not give rise to epistemological problems and do not need to concern us further. For our present purposes, suffice it to say that we depend, usually in an implicit way and unselfconsciously, upon conventionally accepted correlation procedures so that you and I can know—although at times this can be challenging and ambiguous—that we are referring to the same object as perceived by vision, touch, hearing, or another sense. These procedures are, to be sure, not always effective—that is, they do not guarantee that we shall reach a satisfactory agreement that we are in fact referring to the same thing—and they are sometimes unclear and subject to error. Nonetheless, we are able in principle, and are usually able in practice, to correct such ambiguities and errors of correlation.

Despite the occasional inadequacy and fallibility of these procedures, we should be able to recognize that our correlation procedures, like formal rules of transformation between different coordinate systems, establish what we commonly consider to comprise “external perceptual space.” “External perceptual space” is the imagined, conceptualized, or, some would say, the constructed more comprehensive reference frame in terms of which we think of the perspective spaces of individual perceiving observers, or of a single observer, as forming parts. We rely upon the Leibniz boundaries of each individual’s perspective in anticipating that the scope of what is perceived may potentially be extended; we see that other people have similar anticipatory dispositions; and so we imagine, conceptualize, or construct the notion of a more-embracing external perceptual space in connection with which our own individual perspectives supply what we commonly believe is partial information. There is more that needs to be said about this complex matter, which we shall defer until later in this chapter. External perceptual space is the first of the two main varieties of “the external.” We now consider the second.

## 2. *The external ontological world*

By the phrase ‘external ontological world’, I intend to refer to the putatively meaningful commonsense *belief* that objects given in “external perceptual space” continue to exist as they have been perceived, and continue to persist in time during periods when no one perceives them. This belief is commonly

---

this chapter, while the concept of “an individual perspective” and of “the self” will be deferred to a separate chapter.

expressed and maintained (usually unquestioningly) by claiming that the existence of such objects “is independent of our perceiving them.”

This commonsense belief has been advocated even by some of the most theoretically and epistemologically circumspect physicists. Their formulation of this belief is often clearer and less ambiguous than its vaguer commonsense correlate, so let us look at its expression by physicists. We may take Einstein as an example. He repeatedly expressed the following belief in a number of his publications:

Outside there was this large world, which is independent of us human beings and which stands before us as a great, eternal riddle, at least in part accessible to our view and thought.... Physics is an effort to comprehend conceptually that which exists, which is thought to be independent of what is perceived. In this sense one speaks of “physical reality.” (Einstein, 1959/1949b, pp. 4, 80; author’s translation<sup>248</sup>).

The belief in an external world independent of the perceiving subject is the basis of all natural science. (Einstein, 1934, p. 60)

The belief in an external world independent of the percipient subject is the foundation of all science. But since sense-perceptions inform us only indirectly of this external world, of Physical Reality, it is only by speculation that it can become comprehensible to us. (Einstein, 1931, p. 66)

In speaking of his frustration with quantum theory, Einstein wrote:

What does not satisfy me in that theory, from the standpoint of principle, is its attitude towards that which appears to me to be the programmatic aim of all physics; the complete description of any (individual) real situation (as it supposedly exists irrespective of any act of observation or substantiation). (Einstein, 1959/1949a, p. 667)

---

<sup>248</sup> The original German: “*Da gab es draussen diese grosse Welt, die unabhängig von uns Menschen da ist und vor uns steht wie ein grosses, ewiges Rätsel, wenigstens teilweise zugänglich unserem Schauen und Denken.... Die Physik ist eine Bemühung das Seiende als etwas begrifflich zu erfassen, was unabhängig vom Wahrgenommen-Werden gedacht wird. In diesem Sinne spricht man vom ‘Physikalisch-Realen’.*” (Einstein, 1959/1949b, pp. 4, 80).



These claims by Einstein are surely and unmistakably assertions of what Russell called ‘audacious metaphysical theorizing’. We need to look at these claims more closely.

Consider his words: Einstein believed that the task of physics is to understand physical reality “*thought to be* independent of its being observed,” “*as it supposedly exists* irrespective of any act of observation.” The phrases I have italicized turn back the direction of attention from the purported “external world” to *the way we think about it* and *suppose it to be*. By expressing himself in this way, Einstein points to a notion of “external physical reality” that coincides fairly closely with what I mean by “the external ontological world”—namely, *the belief* that objects continue to exist as they have been perceived, and continue to persist in time during periods when no one perceives them. Einstein claimed that we are able to understand such an “external world” “only by speculation.” These are, we may presume, carefully chosen words.

In a later chapter relating to quantum theory ({27}), I examine the concept of the *objectivity* of observations and, by implication, of physical theory itself. Here, what concerns us is the presumptive claim that it *can* make sense to speak of, or to think of, or to speculate about an “external” physical world “as it is thought independently of its being observed,” and whose reality “is independent of any theory.”

We face two questions: Whether it *can* be meaningful to *think* of physical reality “independently of its being observed,” and whether it *can* make sense to claim that such reality “is independent of any theory.” The first makes a claim about what we *can think*, the other about *what is*. The one sets a putatively meaningful task for thought, the other makes a putatively meaningful assertion about reality.

This task for thought falls into the category of “thinking beyond the limits of thought” considered in the previous chapter, and dismissed there as *metalogically* meaningless. We can no more think of physical reality apart from, or while denying, its observability than we can think of what cannot be thought.

Related is the putatively meaningful claim made by Einstein as well as many others that physical reality, as it is studied by theoretical physicists, “is independent of any theory.” Its putative meaning should surely strike us as unverifiable in principle, but its unverifiability is not what concerns us. More fundamentally important, the putative meaning of the claim is projectively self-undermining. The preconditions of reference that would have to be satisfied in order for this claim *possibly* to be meaningful are denied by the very condition set by the claim, that the “reality” purportedly pointed to is “independent of any theory.” To assert such “independence” is to assert the

autonomy of a class of objects with respect to the set of all possible reference frames in terms of which reference to those objects is possible. Physical theories constitute the possible theoretical reference frames in terms of which the objects those theories are capable of studying can be identifyingly referred to.

To assert the independence of the existence of physical objects while withholding the basis for possible reference to such objects is to transgress a metalogical horizon. We shall accordingly call this metalogically self-undermining claim *'the projection of the external world'*. It is a widespread projection which we'll meet in various guises as we proceed. In a later chapter, we shall find that quantum theory, as strictly conceived, intelligently avoids this variety of projection. We will see that it does this by recognizing the physically theoretical impossibility, in principle, of separating the observer and the observer's measuring instruments from what is observed.

### 21.2 The projection of things-in-themselves

[There are] two distinct problems, which it is important to keep separate. First, can we know that objects of sense, or very similar objects, exist at times when we are not perceiving them? Secondly, if this cannot be known, can we know that other objects, inferable from objects of sense but not necessarily resembling them, exist either when we are perceiving the objects of sense or at any other time? This latter problem arises in philosophy as the problem of the "thing in itself," and in science as the problem of matter as assumed in physics.

– Bertrand Russell (1972/1914, pp. 82-83)

I have so far in this chapter discussed two familiar notions associated with "external reality," the first having to do with the "external perceptual world" or the "perceptual space" that is commonly associated with individual sensory perspectives, parts of which perceptual space such perspectives are thought to comprise, and the second relating to the belief in a physical reality independent of its observability. Both of these commonly used notions are expressed in beliefs that putatively claim either (a) that what is perceived is only a part of a larger whole, more of which may potentially be perceived if an individual's sensory perspectives are extended, or (b) that what is perceived is thought to exist independently of possible reference frames in some manner that can be equated with, and perhaps conceptualized as bearing a one-to-one relationship

with, what can be perceived. The first of these beliefs, that of the extendability of perceptual space, we have for now accepted as epistemologically innocent and unproblematic, while the second belief we have rejected as a projective horizon transgression.

In addition to these two beliefs, there is yet a third belief we need to confront: It constitutes another attempt that philosophers have made to reach beyond the reachable, i.e., by positing a realm of *things-in-themselves*. “Things-in-themselves” purportedly populate an in-principle inaccessible world or region of unknowable “objects” that constitute, somehow, the existential or causal basis for the projective “external ontological world” itself.

In short, we appear to be presented, then, with three varieties of alleged reality that are intended to form, so to speak, a group of nested volumes in which perceptual space occupies the center, an “independently existing physical world” occupies the next volume working outward and encompassing the first, and, finally, a third in-principle unknowable volume that is thought to encompass the first two: This last allegedly comprises the outermost region of what has been called ‘things-in-themselves’. This outermost region of things-in-themselves is purported, by those who believe in it, to comprise the reality that lies at the very basis of what we perceive to be real.

As we have recognized in previous chapters, there is a strong human urge to believe and to insist that “there is always more”—that there is no horizon that cannot be exceeded and surpassed, that we can think the unthinkable, that we can express the inexpressible and speak the unspeakable, that we can refer to “what lies beyond the possibility of reference.” This urge is nowhere to be found in a state of such *extremis* as in the claim advanced by both many philosophers and non-philosophers alike that there must lie, beyond the possibility of any experience, a more fundamental, ultimate, and unknowable reality, ever-mysterious and in-principle *epistemologically inaccessible*. The “objects” of this unreachable reality, which make up the domain of “things-in-themselves,” Kant christened ‘*noumena*’, and discussions and controversies among philosophers about such noumena have filled many thousands of argumentative and inconclusive pages ever since. We should be in a position at this point in our study to recognize clearly and explicitly why those many thousands of pages of discussion and controversy have been futile.

Kant was deeply afflicted by the urge to project, to transgress metalogical horizons. He wished to assert the fundamental underlying reality of a realm of objects beyond all human experience, a realm ever-unknowable and ever-unreachable, one which he believed could play the role of the constitutive causal basis of what is commonly believed to be an external perceptual world,

a world commonly judged to exist independently of possible individual experience. “*Noumena*” fit the bill, he thought.

It was an incoherent philosophical wish, as was pointed out by F. H. Jacobi<sup>249</sup> shortly after the 1781 publication of the first edition of Kant’s *Critique*, and then argued again a few years later by Gottlob Ernst Schulze,<sup>250</sup> one of Schopenhauer’s teachers. In abbreviated form, Jacobi and Schulze criticized Kant’s account of noumena as mind-independent objects existing beyond all possible human experience, and intended by Kant to function as a causal basis of our experience of the perceptual world. They argued that his account suffers from an illegitimate extension and application of his own concepts beyond possible human experience and is incoherent. Schulze’s criticism claimed in essence that Kant’s notion of noumena involves the illegitimate application of the concept of causality beyond—translated into our terms—the metalogical preconditions of possible human experience. On the one hand, Kant affirmed that the concept of causality provides a transcendental basis for possible knowledge derived from human experience, but then he illegitimately sought to extend the applicability of that concept to the noumenal realm which, by the terms of its putatively meaningful definition, is in-principle epistemologically unreachable. To do this, Jacobi and Schulze recognized, was not justifiable.

But to do this as Kant did not only lacks justification, it is an attempt to “translocate” a concept which possessed meaning for Kant only in terms of human experience, beyond its metalogical horizon: This is to misapply a concept in an intended context that, in principle, *precludes* that it *can* be applied. To do this is to engage in metalogical projection, to assert what is *not possibly* meaningful.<sup>251</sup>

It would take us too far out of our way to consider the many arguments, some subtle but impressionistic, some technical and analytical, that have been inspired by attempts to conceptualize and talk about this “outermost” region of epistemological unreachability. We will refrain from extending this discussion to philosophers beyond Kant: It ought to stop where it began, with a recognition that the notion of “noumenal things-in-themselves” is inherently projective and, in principle, devoid of sense.

---

<sup>249</sup> Jacobi (1787: see his appendix, “Über den transzendentalen Idealismus,” pp. 209-230).

<sup>250</sup> Schulze (1792).

<sup>251</sup> For a more detailed analysis of similar projective inconsistencies in Kant’s thought, see Bartlett (1970, esp. Sections 2.1, 2.5).

### 21.3 The projection of other minds

I have called ‘the projection of the external world’ the assertion of the autonomy of physical objects, the assertion that is such that, at the same time, a basis for possible reference to such objects is withheld. Similar to this attempted over-stepping of a metalogical horizon is the parallel assertion concerning “other minds.” The belief in the existence of other people’s minds<sup>252</sup> has, like the belief in the “existence of objects independent of our perceiving them,” and like the belief in “things-in-themselves,” served as a rich source of philosophical perplexity and disagreement. Here, too, the method of de-projection is useful.

The belief in “other minds,” like the belief in the external physical world, takes two distinguishable forms: one is epistemologically benign, the other is projective. We look at each in turn.

We have referred to the fact that it is possible to coordinate sensory information that is given in different perspectives, and by this means obtain the notion of a spatial object and of perceptual space. We also saw how the reference frame established by any one sensory modality possesses Leibniz boundaries, which, under suitable circumstances, can lead one to anticipate that those boundaries may be extended. One’s various sensory modalities themselves can, as we experience daily, be correlated so that the individual is able to coordinate information given in different perspectives. Our previous discussion of these correlation procedures was related to the identifiability of *physical objects* within the perceptual space which the correlation of varying sensory perspectives makes possible.

In a similar way, sensory information pertaining to *other people* is experienced in perspectival form, that is, by means of compatible reference frames in terms of which a particular set of objects can be identified—in this case, the set consisting of other people to whom identifying reference is possible. As with our experience of the physical world, sensory information relating to other people is given to us in fits and starts, continuously only for relatively short periods of time, fragmented as attention shifts, broken up by the cycle of sleeping and waking, etc. When such information is correlated—again, in a usually implicit and unselfconscious way—the individual person is able to identify others and to recognize their identities and distinguish among persons. The psychological genesis of the concept of “the other” is not our concern here, but rather what it is possible meaningfully to claim about the existence of “others,” and specifically about the existence of “other minds.”

---

<sup>252</sup> (Also including the minds of non-human animals.)

### 1. *External others*

As in the case of our perspectival experience of physical objects in terms of which we come to understand what I've called 'external perceptual space', so, too, do we come to understand what I shall call 'external others'. A parallel concept of "the external" applies in our recognition that the perspectival experience of other people provides a framework-relative basis in terms of which we are able to correlate sensory experience relating to others and come to identify individual persons. As with external perceptual space, the experience of external others has built into it the anticipatory character that comes to be associated with the factual extendability of the scope of our sensory information that involves other people. Like the notion of external perceptual space, that of external others can be understood simply, without seeking to over-extend the reference frames in terms of which others are identifiable. External others, understood in these minimalist terms, involves no more than the correlational understanding to which the individual is able to come on the basis of the perspectival sensory information available to that individual. The concept of external others, if one were to speak figuratively, forms part of the innermost epistemological "volume" associated with external perceptual space. In these restricted terms, the concept of external others need not give rise to epistemological problems and, like the concept of external perceptual space, will not detain us further.

### 2. *Other minds as independent existents*

In {13.4}, I referred to William Kingdon Clifford's term 'ejects' which he coined to refer to "other minds." We recall his words:

[T]he inferred existence of your feelings, of objective groupings among them similar to those among my feelings, and of a subjective order in many respects analogous to my own, — these inferred existences are in the very act of inference *thrown out* of my consciousness, recognised as outside of it, as *not* being a part of me. I propose, accordingly, to call these inferred existences *ejects*, things thrown out of *my* consciousness, to distinguish them from *objects*, things presented in my consciousness, phenomena.... How this inference is justified, how consciousness can testify to the existence of anything outside of itself, I do not pretend to say.... (Clifford, 1878, p. 58 )

By extension, making an interpretive step to which Clifford might not agree were he available to consult, I believe his notion of “ejects/other minds” may have been intended by him to express the putative meaning of “other minds that are independent of our experience of them.” “Other minds” are, in his view, “inferred existences” that are “*thrown out of my consciousness*”—as distinguished from objects or things that are given in my consciousness, as are, in our terms here, physical objects and the external perceptual world.

I quote Clifford in this context, not to because it is important to our present purposes whether the view I’ve hypothetically attributed to him is correct, but rather to illustrate a conception of “other minds” that seeks to go beyond the metalogical horizon of possible reference, and, by so doing, attribute a putative reality to such “other minds” in much the same way as we saw in connection with the projection of the external world.

The putatively meaningful notion that “other minds exist independently of our perceiving them,” as in the case of the putatively meaningful notion of the external ontological world, attempts to assert the independent existence of “other minds” at the same time as a basis for possible reference to such objects is withheld, thereby transgressing a horizon of possible reference and meaning. We shall accordingly call this metalogically self-undermining claim ‘*the projection of other minds*’. To claim that “other minds” putatively exist independently of one’s possible experiential framework is to fall victim to this projection. It is metalogically self-undermining and therefore devoid of meaning.

Russell recognized that the epistemologically naive belief in what I’ve called ‘the external ontological world’ is on a par with the equally naive belief in putatively independently existing other minds: Each of these beliefs, Russell noted, “is felt to demand logical justification” (Russell, 1972/1914, p. 79). But, in each instance, neither belief is capable, in principle, of justification because, in each instance, neither purportedly meaningful belief can, in principle, make sense because each undermines its own referential basis. To claim that either belief is meaningful is to fall victim to a metalogical sleight-of-hand that brings about an apparently meaningful delusion of reference.

### 3. *Other minds as things-in-themselves*

As was the case with physical objects, some philosophers have wished to trespass even further beyond the metalogical horizon of sensory experience by means of a projective attribution of autonomous existence to other minds. This further projective step appears to be taken as a result of two varieties of awareness. There is, on the one hand, the awareness that other people, as perceived perspectively, possess Leibniz boundaries that encourage one to

anticipate that additional sensory information about them may be forthcoming with the passage of time and changing conditions of perception. On the other hand, there is the framework-relative awareness of the individual percipient, whose projective predisposition inclines him or her to claim that, just as he or she perceives, feels, experiences, and thinks, so putatively does the other person—but in a manner that is believed to possess an epistemologically inaccessible “noumenal existence” independently of any possible reference frame to which the individual percipient has access. This is the ultimate degree of epistemological inaccessibility that we encountered in connection with things-in-themselves.

We can and do come to recognize, and sometimes also understand and empathize with, the perceived behavior and expressed feelings of others, and we do this, as we do in connection with physical objects, through perspectival sensory experience and our abilities to correlate information given in different perspectives. However, when a putatively meaningful additional step is taken, attempting to over-step the horizon of the individual’s possible experiential reference frame, a self-undermining projection results. It is no more possible meaningfully to claim that “other minds are independently existing things-in-themselves” than it was to make the same claim about physical objects. To distinguish this projection from the “projection of other minds,” we may appropriately call it ‘the projection of other minds as things-in-themselves’.

#### **21.4 The de-projective understanding of the external world and of other minds: The need to relinquish these projections**

If we adhere to the strongly compelling standards of rationality defined in {17.3},<sup>253</sup> and if we accept the metalogical incoherence of the projections of the external world and of other minds, then how are we to understand the commonly used and widely relied upon concepts of the external world and other minds? We must recognize that, although both concepts are projective, there is, for each, a de-projected meaning that can be salvaged and accepted as epistemologically coherent.

In both cases, we recognize that the two classes of objects of reference, physical objects and other minds, cannot in any way be asserted to be autonomous of reference frames in terms of which they are identifiable. But in rejecting these projections, we do not thereby affirm their negations: It is

---

<sup>253</sup> We recall that, in abbreviated form, the two fundamental rules of rational acceptability are: (i) a given frame of reference must be free of projection, systemically coherent, and must have the capacity to refer to a range of objects of reference; (ii) a rational agent must be such that his or her rationality and conviction are fused.



equally projective to claim that physical objects and other minds “do not possess autonomous existence” and “depend for their existence” upon frameworks of identification.

When de-projected, what lies at the core of this commonplace intention to assert such putative independent existence, whether of physical objects or of other minds? In both cases, it is my impression that non-philosophers, who have not become entangled in age-old mystifications created by philosophers, intend, on a fundamental, commonsense level, to express their underlying beliefs in the inferred *continuity* and perceived *regularity* both of physical objects and of other minds, despite the acknowledged fact that sensory experience is variously fragmentary and interrupted. Despite the fact that sensory experience is given in discontinuous perspectival form, sensory experience of physical objects and of other minds provides a basis for the inferred continuity of many of the objects we perceive, as well as the inferred continuity of people we experience. We also perceive that our sensory experience is characterized by a certain degree of regularity: We find upon awakening from a night’s sleep, for example, that our sensory experience today has much in common with what we remember having perceived yesterday; interrupted periods of experience show varying degrees of regularity in this sense.

The belief in continuity is the common and pervasive belief that within “external perceptual space” both physical objects and other persons are characterized by Leibniz boundaries which, when circumstances are right, may potentially be extended; this is the belief that, both with respect to physical objects and other minds, the scope of what we perceive may potentially be expanded. We come to form a habitual anticipatory disposition that is expressed in our beliefs in the general continuity and regularity of sensory experience.

To infer continuity as a function of Leibniz boundaries is metalogically unobjectionable; nothing in such an inference is inherently projective. However, there is among many philosophers and non-philosophers alike a projective disposition to assert that, corresponding to the continuity which we infer, there is an independently existing reality that forms the basis for such continuity, a continuity that is claimed and believed to be autonomous of frames of references relative to which such reality has possible meaning. We shall call this the ‘*projection of general continuity*’. To claim this is metalogically self-undermining and, rationally, that claim must be rejected.

Although many philosophers *wish* for more than this, “more than this” is not, in principle, forthcoming for the rational mind. Whether based purely upon sensory information or upon conceptual thought, the framework relativity of our experience does not provide us with a possible means to know or

meaningfully to say more than what we *can* know or *can* say in framework-relative terms.

As we have already had a number of occasions to observe in this study, the human predisposition to engage in delusional metalogical projection is widespread and seductive. It should come as no surprise that when we perceive regularity in the external perceptual world, and when we infer continuity of objects of reference despite the discontinuity of observation, the same widespread projective predisposition will frequently be expressed. That predisposition takes the form of projective beliefs in a putatively autonomous existing physical world and in the putatively meaningful “independent existence,” and sometimes even in the “noumenal existence,” of other minds.

De-projected, the concepts of an external physical world and of other minds reduce to the meaningful beliefs I have described. To wish, and to assert, that “there is more than this,” is to seek to transgress the metalogical horizons of the frameworks of reference in terms of which physical reality and other minds may, in principle, be said to be and have meaning.

### 21.5 The projections of realism and idealism

Let us begin with a simplified representation of these two infamous sources of philosophical controversy. Following this minimal schematization we’ll turn to examine the general theses of realism and idealism as they have been propounded by philosophers.

In the following example,<sup>254</sup> we consider realism and idealism solely with respect to physical events (hereafter simply called ‘events’). Assume that the position that I shall call ‘event-realism’ with respect to such events is characterized by the claim:

Events exist that I not only do not perceive, but these events exist independently of my perception. (Claim A)

Further, assume that the position of “event-idealism” with respect to physical events is characterized by the contrary claim:

There do not exist events that I do not perceive, and the existence of events is dependent upon my perception. (Claim ~A)

---

<sup>254</sup> This example is in part based on Bartlett (1983a, pp. 86-88).

Let the following set of propositions, taken together, express the simplified thesis of event-realism; the purpose here is to render explicit the obduracy that knows no limits which characterizes the realist's position:

1. There is a frame of reference  $F$  that establishes the only basis in terms of which reference is possible to events. In  $F$  the concept of an event has its standard, commonsense meaning of a physical occurrence taking place at a certain place and time.
2.  $F$  provides a basis for reference to a set of events  $E$ .
3. There are events not included in  $E$ ; call this set ' $\bar{E}$ '.
4. Events in  $\bar{E}$  are events in the same sense of those in  $E$ .
5.  $F$  does not provide a basis for reference to  $\bar{E}$  or events in  $\bar{E}$ .
6. There does not exist any other frame of reference by means of which reference is possible to  $\bar{E}$  or to events in  $\bar{E}$ .

We see immediately that the above set of propositions is patently inconsistent: Propositions 3.–6. refer to  $\bar{E}$  or events in  $\bar{E}$ , while proposition 6. in conjunction with 5. *precludes* a basis for such reference. In addition, 3. and 4. refer to events not in  $E$ , yet 1. asserts that the only frame of reference capable of referring to events is  $F$ , which, according to 5., does not permit reference to events not in  $E$ . Yet the thesis of realism requires 6. because realism intends to assert the *autonomy* of physical events in relation to  $F$ , the only available frame of reference. The important point here is this: Were some additional frame(s) of reference to be made available—e.g.,  $F'$ , enabling reference to events in  $\bar{E}$ —the intention of realism will still be to insist that there *exist* events extending beyond those to which  $F$  *conjoined with*  $F'$  permit reference; however we may generously add further frames of reference is irrelevant to realism's relentless insistence that events exist autonomously of any given reference frame. The purpose of proposition 6. is to save realism from this potential regress by stipulating that there is no other frame of reference by means of which reference is possible to events other than those in  $E$ . But even given the stipulation expressed by 6., realism will stubbornly persist in insisting, in its motivation to reach beyond its grasp, that other *events*—beyond whatever reference frames are provided—*do* exist.

Taken together, the set of propositions 1.–6. is of course internally inconsistent. Any attempt to construe event-realism in any sense that is derivable or reducible to any set of propositions equivalent to 1.–6. is clearly inconsistent.

Let the following set of propositions, taken together, similarly express the simplified thesis of event-idealism:

- 1'. There is a frame of reference  $F$  that establishes the only basis in terms of which reference is possible to events. In  $F$ , the concept of an event has its standard, commonsense meaning of a physical occurrence taking place at a certain place and time.
- 2'.  $F$  provides a basis for reference to a set of events  $E$ .
- 3'. There are *no* events not included in  $E$ . The set  $\overline{E}$  is empty.
- 4'. That the set  $\overline{E}$  is empty expresses the idealist claim that  $\sim F \supset \overline{E}$ : i.e., without framework  $F$ , there *are* no events.
- 5'.  $F$  does not provide a basis for reference to  $\overline{E}$  or to possible members of  $\overline{E}$ .
- 6'. There does not exist any other frame of reference to refer to  $\overline{E}$  or to possible members of  $\overline{E}$ .

The above set of propositions, like 1.–6. earlier, is internally inconsistent. The proof is similar to the case of event-realism, substituting 1'–6' for 1–6.

We now turn to examine the more general philosophical positions of realism and idealism.

### 1. *Realism*

Realism as a philosophical position continues after millennia to attract the attention and contention of philosophers. The position can be expressed in different ways, depending upon what sorts of “objects” the realist claims to be “real.” But whether the objects whose reality the realist wishes to assert are claimed to comprise ideas, forms, universals, truths, physical objects, or other minds, the central putatively meaningful belief of realists is that these objects possess an existence independent of the mind, independent of the perceiver, independent of the observer and of the observer’s means of observation. To assert the “independence” demanded by realism, as we shall understand it in its strongest form, is to assert the autonomy of the existence of one or more classes of objects with respect to the set of all possible reference frames in terms of which reference to those objects is possible.

In {13.4}, I quoted Josiah Royce’s objection to realism because it exemplifies an apparent step closer to the concept of projection developed in this study. We recall that he argued that realism denies its own presuppositions by

attempting “in vain to articulate” a meaning that its own presuppositions render impossible. For Royce, “the realm of valid or possible experience” comprises what I’ve previously called a ‘closed system’. In attempting to claim that objects “exist independently” of “the realm of valid or possible experience,” Royce concluded that realism involves “a hopeless contradiction” that it is “wholly inconsistent” and “meaningless.”

We saw in {4}, {10}, and {15} that a general system of reference comprises a closed system despite its absence of determinate boundaries. We observed that in relation to and solely from the standpoint of such a system, attempts to refer “outside” of that framework will, of necessity, presuppose the referential means supplied by that framework, and, as a consequence, such attempts to reach beyond that framework’s metalogical horizon undermine themselves: they undermine their own possibility of reference and meaning. We have encountered a number of such projective attempts so far in this study. Realism is yet another and very unmistakable attempt at horizon transgression.

Although a given frame of reference may possess Leibniz boundaries that in principle can be extended, yet the referential field it establishes has no delimited boundaries. But as we’ve seen, it does possess what I’ve called ‘reactive boundaries’ that are made evident by attempts to trespass beyond those boundaries; these reactive boundaries disclose that reference frame’s metalogical horizon. The realist philosopher who claims that “objects exist independently of our means of apprehending them” attempts to transgress such horizons and falls victim to projective delusion. Whether those means are purely mental, sensory, or rely upon instrumentation, the position of realism is metalogically self-undermining and meaningless. As shown in {18}, the very identity of any object is indissolubly linked to the reference frame(s) that provide a basis for its identifiability: that link provides information that expresses in what sense the object *is*. In realism we encounter once again a manifestation of the human psychological and cognitive urge or compulsion to engage in metalogically projective thinking, to over-extend claims beyond the frameworks of their possible meaningfulness.

## 2. *Idealism*

The philosophical position of idealism, as the reader is now likely to foresee, is no less projective than its contrary cousin, realism. Like realism, philosophical idealism has been formulated in different ways, depending upon the kinds of objects the idealist believes are mind-dependent. The central contention of idealism is that the kinds of objects at issue for the idealist putatively

*depend* for their existence upon the mind, consciousness, or experience of the idealist—that is, they purportedly have no separate, independent existence.

Where the rationale for realism may be found in the general uniformity and regularity of sensory experience, the motivation for idealism may, we conjecture, be found in the fact that experienced objects vary in their experienced properties in correlation with variations in the physical or mental state of the idealist. For example, when the idealist closes his or her eyes, visual objects are no longer experienced—although there may be exceptions, of course, in the form of after-images, visual hallucinations, etc. Sensory correlations like this may suggest to the idealist a rationale for generalizing that which he or she takes to be validly inferred from such experiences—for example, that physical objects putatively depend for their existence on the idealist’s reference frame.

Other sources of motivation that may lead to idealism include the putative experience of “creative invention” (see {19.1}), which is believed to lend support to the notion that the mental activity of the idealist brings about solutions to problems (those solutions “would not exist were it not for such mental activity”). The result from these and similar experiences has very likely inspired some philosophers to propound idealism in the generalized claim that “to be is to be perceived.”

Whatever its motivation, the idealist’s claim that objects are “mind-dependent”—possessing an existence that purportedly depends upon the consciousness or experience of the idealist, but allegedly having no separate, independent existence—is evidently projective, as the reader should by now immediately recognize. The defining conditions established by the idealist’s claim rule out the possibility of any framework of reference in terms of which preconditions necessary for reference to such a putative dependency relation can, in principle, be satisfied. Satisfying the necessary conditions for such possible reference is barred by the terms of the idealist’s position, which therefore undermines its possible meaning.

...

In this chapter, we’ve seen how the putatively meaningful concepts of an external world, of things-in-themselves, of other minds, as well as the positions of realism and idealism, all, and in similar ways, transgress their metalogical horizons. All make claims that attempt to reach beyond the metalogical horizon of the frame(s) of reference that each presupposes. All undermine their own possibility of meaning.

It can of course be intellectually and psychologically challenging to relin-

quish projective concepts and ways of thinking that feel intuitively persuasive and commonsensical. The task of developing a thoroughly de-projective understanding that complies with the demands of framework relativity requires disciplined analysis, often hard work, and a willingness and a will to reject beliefs that fail to meet rigorous standards of rationality. These conceptual and psychological skills are not frequently to be found combined in one person. Readers interested in basis for the latter claim and in a more detailed description of the requisite skills are referred to Appendix II, "Epistemological Intelligence."

## 22

### The Projections of Time, Space, and Space-time

The concepts of time and space have entered most of the discussions in previous chapters only indirectly and have often been employed with their commonly accepted meanings and without reflective analysis. In this chapter, we examine various ways in which the concepts of time and space are employed frequently and habitually in self-undermining projective ways.<sup>255</sup>

#### 22.1 The temporal basis of reference

In {10.1}, it was noted that identifying reference to an object, of no matter what kind, is only possible when that object of reference is in some way fixed or defined within its presupposed framework so that *re-identification* of that object is possible. Why should this be the case? The re-identifiability of an object allows us to recognize that one and the same object is in view in a plurality of referential contexts. Were we unable to re-identify an object, reference to that object would not be possible in a manner so that we could make claims about it, communicate about it, or know what we are talking about. The re-identifiability of objects of reference requires that a reference frame permit *temporally successive*, reiterated reference, so that, as expressed earlier in this study, retrospective second-order references are possible to past references. This is another way of saying that the possibility of reflection presupposes ways of retaining temporally earlier states. We can, for example, only think and talk about that concerning which we retain some memory or other record. Whether human memory is involved, or reliance upon a recording device, or dependence upon evidence or other signs that we consider to comprise artifacts of the past, the identifiability of objects of reference requires that temporally successive, iterable references to the same object be possible.

---

<sup>255</sup> While the present study examines projections of time and space from the standpoint of the metalogic of reference, Bartlett (1970, Section 2.1) provides an analysis of a number of corresponding projections from the standpoint of phenomenology.



In {12.1}, we also noted that the capacity to recognize error, and specifically in this study the capacity to recognize metalogical projections, fundamentally requires that a frame of reference provide a basis for the re-identifiability of the objects to which it can refer.

Understood in terms of the functional relativity of any set of objects of reference with respect to the reference frame(s) in terms of which they can be identified, identifying reference has an important and fundamental temporal basis. This is not by any means to suggest that every object of identification must specify a time, or a time coordinate, but rather that identifiability itself presupposes the possibility of reiterated reference, and reiterated reference requires that retrospective, second-order, references be possible to past references.

What is in view here is the serial nature of successive, reiterable reference, necessary in order that “the same” object of reference can be recognized by an individual observer at subsequent times, or by subsequent observers. The conception of “subsequent times” is a function of and derives its meaning from this very successive reiterability of reference. The temporal basis of reference is, in the sense developed in this study, metalogical in that it comprises a precondition of identifying reference. The temporal basis of reference is logically fundamental to the capacity of a reference frame to identify spatial objects, the capacity of which, for its possibility, presupposes re-identifiability and hence a temporal basis. It is therefore appropriate that we first consider projections involving time.

In describing the temporal basis of reference, we are not in any sense speaking about “time” as an idea, but about a framework’s constitutive structure that provides for the capacity for sequential or serial retention of earlier states. It is important that it should be clear to the reader that the metalogical, conceptually fundamental, temporal basis of reference is to be distinguished from the specification of time coordinates of objects of reference.

In {15}, we saw that when a theory is subjected to de-projective analysis, an examination is made of how that theory functions as a system of coordination capable of interrelating the following: the set of its putative objects of reference, a formal or other system that provides a presupposed context of reference, and perhaps explicit reference to a time, location, and/or relationship to one or more observers or recording devices. Through such an analysis, the parameters of constraint are made explicit that define the range of reference that is possible from the standpoint of that theory. Undergirding the possibility of such an analysis is the metalogical temporal basis of reference described in the previous paragraphs. Without this conceptually fundamental temporal capacity, no reference frame is capable of identifying reference.

## 22.2 Projections of time

We turn now to consider a variety of projections that have their basis in time-order. Several such projections will be described.

### 1. *Projections of the past*

One of the most common, pervasive, putatively meaningful claims that many philosophers and non-philosophers alike make concerning past events is the claim expressing the belief that those events possess an autonomy of their own and can be presumed to “have existed” in some sense independently of what we consider to comprise the present. Whenever purportedly meaningful claims like this are made, a context of reference is of course presupposed in terms of which events are identified as past. Such events are perceived as past through any of the means available that provide us with ways of retaining information about events, such as human memory, recording devices, physical evidence, etc. Each such context of reference makes use of a reference frame in terms of which recalled information is identifiable, whether specifically, vaguely, or in some rule-determined way.

Reference to past events is subject to the same framework-relative constraints studied in earlier chapters. To attempt to regard retained or recorded information that defines past events as possessing, in any meaningful sense, an independence with respect to the frameworks presumed by reference to those events, is immediately recognizable as metalogically projective. Such an attempt to “sever” the relativity of what is recalled, remembered, or otherwise recorded, from reference frames that must be presupposed for such recollection, memory, or recording to be possible, is metalogically self-undermining and meaningless. We shall call such projective claims that seek variously to assert the “autonomy” or “independence” of past events ‘*projections of the past*’.

### 2. *Projections of time-flow*

A second temporal projection is involved in the putatively meaningful notion that what we regard as comprising “the present” in some way “becomes past.” This projection is expressed in the purportedly meaningful notion that this “becoming past” involves “a process in which the present becomes past.” From this, the putatively meaningful claim often is made that there exists a process of “passage of time” or “time-flow.”

De-projected, these notions and the beliefs to which they lead are based on a set of familiar and unproblematic facts that are misconceived: Reference to past events is possible and is meaningful only in intrinsic relation to

frameworks that have available certain means to retain information in terms of which we come to define and understand what is past. We observe the temporal succession of events, paralleling our observation of the succession of visual perspectives of a physical object. The recognition that events are observable in temporal succession is itself unobjectionable from the standpoint of de-projective analysis. However, as was the case in connection with the projective commonsense notion of a physical reality supposedly independent of frames of reference in terms of which such reality is understood, observations of events given in temporal succession—given, to speak metaphorically, in “temporal perspectives”—provides no possible basis for asserting that “time itself is passing” or that “there exists a flow of time.”

The belief that “time flows” or that “the present becomes past” would, in order in principle to be meaningful, require an altogether different and in principle inaccessible frame of reference, one that would “stand apart” from the serial nature of temporally successive reference. Such a belief, to be possibly meaningful, would require the breaching of a metalogical horizon which, in principle, cannot be transgressed. We shall call this variety of temporal claim ‘*projections of time-flow*’.

### 3. *Projections of the future*

As in connection with past events and the widespread human tendency to embrace the projective belief in an alleged framework-independent past, there is a similar inclination to believe in a putatively meaningful, framework-independent future. We use a wide variety of means to refer to future events, situations, outcomes of experiments, etc. We anticipate that certain occurrences are likely to happen: We forecast the weather, calculate probabilities, estimate the likelihood of nuclear war, etc. Whether an individual’s personal experiential frame of reference is in view, in relation to which he or she has come to form certain expectations, or whether predictions are made in accordance with a physical theory as it applies to a given set of conditions, or whether the specific outcome of a chemical reaction has come, in a law-like manner, to be expected, we habitually—by experience, or by training, or by virtue of the theories we endorse—develop beliefs and make anticipatory or predictive claims with which we associate varying degrees of certainty.

However, we often lose sight of the essential connection between our anticipations or predictions relating to events and situations that have not yet occurred, and the frames of reference which these anticipations and predictions themselves presuppose. When we do not acknowledge this necessary connection and develop the belief in an alleged temporally autonomous future—which seems to many of us apparently meaningful, commonsensical,

and intuitive—we fall victims to a class of temporal projections which it is appropriate to call ‘*projections of the future*’.

Much like the anticipations that we habitually form as we perceive three-dimensional physical objects from various perspectives, so do we come to form patterns of anticipation that express our beliefs in what is most likely to occur in many situations. When these beliefs become strong, having been confirmed and re-confirmed many times, the projective belief can become firmly established that there exists a future whose status, as essentially predictable or probable to a high degree, is thought purportedly to be independent of the frameworks of reference that we rely upon in the very formulation of future expectations. As physicist P. W. Bridgman (1936, p. 32) expressed this, “We usually think of the future as stretching before us, and ourselves going to meet it. Who knows how much this picture may be responsible for the feeling ... that the future has ‘existence’ and is essentially predictable...”.

It should quickly be clear to any reader who has come this far in this study that, in principle, no reference frame is possible in terms of which such an “independently existing future” is identifiable. There is no possible way in which this putatively meaningful notion and the group of words employed to express it can refer. Belief in the purportedly meaningful “independent existence of the (or a) future” is projectively self-undermining.

We shall find in a later chapter that this projective belief in an “independently existing future” has served as a stumbling block for those physicists and philosophers of science who insist upon the reality of events which they believe possess framework-independence.

#### 4. *Projections of absolute time*

When we take collectively into account the previous results concerning projections of the past, of time-flow, and of the future, it immediately follows that the notion of an “absolute, independently existing continuum of time” must also be dismissed as inherently projective. Once the putatively meaningful notions of an “independently existing past,” of “the flow of time,” and of an “independently existing future” are rejected and relinquished as metalogically meaningless, we find that there can be no possible meaning associated with the notion of an “absolute, independently existing continuum of time.” It is only the framework-relative succession of events that provides a basis for the de-projected meaning of the recognition and measurement of change of state, in terms of which temporal coordinates and the concept of temporal intervals have their conceptual and observational basis.

### 5. *Projections of temporal constitutive subjective activity*

A fourth variety of temporal projection may be mentioned here. It concerns a putatively meaningful notion that has made its way into some philosophers' attempts to understand the concept of time and of temporal succession. The accounts given by Kant and Husserl may serve as examples. —Once again, as previously in this study, my concern here is not refined and carefully qualified textual explication, but an abbreviated characterization of a philosophical position for the purpose of illustration.

In Kant's *Critique*,<sup>256</sup> particularly in Edition A, in describing the Forms of Sensibility, Kant maintains that an active subjective imposition of the form of time is made upon a chaotic manifold of representations. In my reading, what he appears to have wished to claim is that the time-order of human experience is the "result of a constituting activity," an active imposition of time-ordering by the mind upon what otherwise is chaotic and without the temporal structure characteristic of and essential to experience. Such temporal structuring is indispensable, in his view, to the intelligibility of experience and to the possibility of knowledge that we are capable of gaining from it.

In a similar way, Husserl in several works<sup>257</sup> advocated the notion of "acts that are constitutive of time." In Kantian fashion, he wished also to advocate the putatively meaningful notion that subjective mental activity is responsible for the temporal constitution of experience.

Neither of these claims, whether by Kant or by Husserl, can—in principle—have possible meaning. Were the human mind to be so structured as they describe, no humanly accessible reference frame would be possible to permit reference to the putatively drawn distinction between human experience, as temporally structured, and "chaotic representations" that allegedly form the raw material for the "constitutive activity" of the mind. The attempt to make this claim transgresses the metalogical horizon established—in accordance with their own views—by the human mind and its experience. We shall therefore call this variety of temporal projection '*projections of temporal constitutive subjective activity*'.

## 22.3 The nature of spatial order

In the previous chapter, "external perceptual space" was understood as the imagined or conceptualized comprehensive reference frame in terms of which

<sup>256</sup> Kant (1965/1929, A 95, 97-106, 109, 118ff, B 160, 211f).

<sup>257</sup> Husserl (1964/1905-1910, §15, p. 53, §16, p. 55), (1965/1929, §61, p. 222), (1962/1936, pp. 176, 178).

the perspective spaces of individual observers are thought to form parts. We saw that, despite the fact that sensory experience of physical objects is largely given during periods of discontinuous observation, it supplies a basis for the inferred or imagined continuity of many perceived objects and is often characterized by a certain degree of regularity. In describing external physical space in this initially broad way, we have yet to make clear the fundamental nature of spatial order.

This chapter began with a discussion of the temporal basis of reference. We now need to extend this discussion as it applies to spatial objects. We begin having recognized that a precondition of identifying reference is temporal, while the identifiability of spatial objects presupposes their re-identifiability and hence possesses a temporal basis.

Both temporal and spatial forms of order are serial in nature. We recall that ‘serial’ here refers to the nature of successive, reiterable reference, providing a basis so that one or more observers may recognize “the same” object. The metalogically fundamental form of serial order we have called ‘temporal’. Spatial order requires temporal serial order, which we recognize as we perceive physical objects in serially given perspectives. The imagined or inferred continuity of a physical object is itself serially based. A spatial object of one dimension, a line, is considered to be continuous by virtue of its serial ordering of the successive points that constitute it. This is equally true of a two-dimensional surface and of a three-dimensional solid. The serial ordering nature of temporal succession is, as it were, “translated” to the dimensions we associate with spatial order.

From the level of maximum theoretical generality, the serial order of events to which we refer as past, present, or future is, in terms of its constitutive structure, no different in kind as we find in the serial order of physical objects to which we refer as having spatial dimensions. So understood, spatial order embodies the serial nature of temporal order where the dimensions of serial order have increased beyond the single dimension of temporal succession. We may, then, regard spatial order to consist in no more than a “translation” of the serial, sequential nature of temporal order to objects given in successive perspectives. Alternatively expressed, a reference frame that makes it possible to identify physical objects must be temporally based, while the constitutive structure of the physical objects to which it permits reference is both serial in nature and given serially in perspectival form.

To illustrate this:<sup>258</sup> Consider a spatial object  $S$  that is understood in terms of partially overlapping sets of perspectives given over an interval of time  $t_1$  to  $t_3$ :  $\{s_1, s_2, s_3\}$  is given at time  $t_1$ , while  $\{s_2, s_3, s_4\}$  and  $\{s_3, s_4, s_5\}$  are given

---

<sup>258</sup> This illustration is employed, in part, in Bartlett (1970, Section 2.1).

respectively at  $t_2$  and  $t_3$ , where  $t_1$  is earlier than  $t_2$ ,  $t_2$  is earlier than  $t_3$ , and  $t_1$ ,  $t_2$ , and  $t_3$  are identified from the standpoint of a frame of reference which provides a basis for reference to what we consider to comprise the present time. The set  $s$  of variations  $\{s_1, s_2, s_3, s_4, s_5\}$  relative to the relations expressed between  $\{s_1, s_2, s_3\}$ ,  $\{s_2, s_3, s_4\}$ , and  $\{s_3, s_4, s_5\}$ , is then said to be constitutive of a spatial object  $S$  given in successive perspectives at  $t_1$ ,  $t_2$ , and  $t_3$ . Since, as we have seen, spatial objects typically have Leibniz boundaries, the possibility of extending the set  $s$  may also be constitutive of the particular spatial object  $S$ .

From this illustration we see that the serial order of spatial objects is essentially temporal: The identifiability of physical objects is temporally based, while physical objects are essentially relative to reference frames that permit a retentive record of their serially given past states.<sup>259</sup>

## 22.4 Projections of space

We turn now to consider a variety of projections that have their basis in spatial order.

### 1. *Projections of absolute space*

Isaac Newton, in his historically influential 1606 treatise, *Mathematical principles of natural philosophy and the system of the world*, made the claim that “space” is independent of the existence of material bodies, which together make up the physical universe.<sup>260</sup> In the view Newton endorsed, space is a continuum that may either be empty or be occupied by physical bodies. This continuum purportedly provides a system of reference in terms of which the position of any material body can in principle be discretely located, that is, can be assigned coordinates that specify in a completely determinate manner its location in the independently existing continuum.

The commonsense conception of space that persists to this day is fundamentally Newtonian: What I’ve called ‘external perceptual space’ is commonly thought to comprise what figuratively might be called a ‘container’

---

<sup>259</sup> Among philosophers who have to varying degrees recognized the basis of spatial order in the serial nature of temporal order, Whitehead (1961/1922, p. 301) observed “order in space is merely the reflexion into space of one time-system of the time-orders of alternative time-systems,” while Husserl (1913-1921/1900-1901, *Untersuchung III*, supp. notes, p. 256) claimed that “objects are [unities] as a function of time,” and he later noted that “the constitution of the spatial object ... presupposes ... that of time” (Husserl, 1928, §43).

<sup>260</sup> Newton (1934/1606, pp. 6-12, 639-644). This claim is also made in Newton (1952/1704, Book III, *Queries* 18ff).

within which the usually temporally discontinuous perspective spaces of individual perceiving observers, or of a single observer, are thought to form parts. Paralleling Newton's notion of absolute space, the commonsense view is that this "container" serves as a volume which is independent of the material bodies that fill it and in terms of which observed spatial objects can be located.

In the previous chapter, we saw how the notion of external perceptual space frequently leads to the allegedly meaningful belief that objects given in this space continue to exist as they have been perceived, and continue to persist in time during periods when no one perceives them. It is a belief that is frequently expressed in the claim that the existence of such objects "is independent" of the perception of observers. Earlier in this chapter, we saw that such a belief metalogically entails a self-undermining projection, which I've called the 'projection of the external world', while for some epistemological realists this belief takes its extreme form in the projection of things-in-themselves.

In Newton's notion of absolute space we find much the same projection expressed. When Newton endorsed the notion of absolute space, we see, considered from the standpoint of the metalogic of reference, that he implicitly must affirm three things: (a) that identifying reference to material bodies requires an appropriate coordinate system in terms of which reference to spatial objects is possible; (b) that there exists a continuum of space which he claimed serves as that coordinate system; and (c) that this continuum is autonomous of whatever material bodies may, or may not, exist in that continuum: its independence is absolute.

Newton's self-undermining projection of course occurs in affirming (c): For reference to be possible to physical objects, an appropriate framework of reference must of course be presupposed: As we have seen, it must permit temporally successive and retentive reference. Further, such a framework of reference permitting reference to spatial objects is, as we've noted, nearly always temporally discontinuous—observations of physical objects are seldom continuous observations, but instead are given in perspectival form. Objects which we perceive as spatial and as situated in an imagined or conceptualized comprehensive "external perceptual space" are essentially relative to the perspective spaces of observers. This relativity of spatial objects to observer-based spatial perspectives is constitutive of—i.e., is a precondition of—what we understand by "material bodies." To attempt putatively to represent an object of reference as spatial while the preconditions for reference to such objects cannot be satisfied, results in self-undermining projection. The preconditions for reference to an absolute continuum of space considered to be independently of the very preconditions of reference



that allow us to understand spatial order is of course impossible. This is the consequence to which Newton's concept of absolute space inevitably leads. We shall accordingly call such projections '*projections of absolute space*'.

Kant's Newtonian notion of absolute space was nearly as influential among philosophers, at least for a time, as was Newton's among later physicists. In his 1770 *Inaugural Dissertation*,<sup>261</sup> Kant argued that the relation of parts of space to one another presupposes a region in relation to which they are possess this order. This relation ultimately consists, he claimed, not in the relation of one spatial thing to another, but in the relation of the system of their positions to what he called 'absolute world-space'. This "world-space," Kant claimed like Newton, is independent of the existence of all matter, and is the basis for the compositeness of matter. He maintained that the reality of differences between individual physical objects is made possible only due to their relation to this "absolute world-space."

Having reached the conclusion that Newton's notion of absolute space is metalogically projective, there is no need for us to retrace the same steps, *mutatis mutandis*, as they apply to Kant here. If "world-space" comprises—translating Kant's notion into our terms—the necessary reference frame for possible identifying reference to spatial bodies, its claimed putatively meaningful independent spatial existence requires the possibility of reference to that to which reference is impossible. Reference to a "world space" autonomous of all matter is not only physically impossible, but also conceptually impossible when such purported reference would undermine the preconditions of reference to spatial order. Kant's "absolute world-space" and Newton's "absolute space" are both metalogically projective and equally meaningless; both attempt to sever the framework-relative ties of spatial objects to the referential frameworks that permit reference to them.

Both Newton's putatively meaningful notion of "absolute space" along and Kant's notion of "absolute world-space" lead, when de-projected, to a concept of space that is a function of the identity of physical objects of reference—whose identity and ontology are essentially relative to the parameters of constraint of reference frames in terms of which they are identifiable.

We have of course reached this conclusion by the means provided by the metalogic of reference, i.e., by a means that is purely epistemological, not a means generally employed by theoretical physics. And yet we shall find in subsequent chapters ({26–28}) that this conclusion directly re-affirms the framework-relative results of both relativity physics and quantum theory.

In connection with relativity physics, it is interesting to note that Einstein, as a consequence of the general theory of relativity, concluded that the

---

<sup>261</sup> Kant, Immanuel (1929/1770, pp. 19-29).

concepts both of time and of space must be relinquished as having autonomous existence. He set the requirement that the laws of physics are to be expressed in general covariant form, that is, in a generalized manner that can be transformed to any frame of reference without loss of validity irrespective of the physical conditions of whatever reference frame may be employed. To require this is, in Einstein's words, to require that physical laws express "universal covariance which *deprives space and time of the last vestige of physical objectivity*" (Einstein, 1916a, p. 776, italics added).<sup>262</sup> As a consequence, Bertrand Russell observed "... space and time have ceased to be, for relativity physics, part of the bare bones of the world, and are now admitted to be constructions" (Russell, 1972/1914, p. 109).

## 2. *Projections of spatial constitutive subjective activity*

Much as he did in connection with the allegedly meaningful temporal constitutive subjective activity of the mind, Kant believed that the mind performs an active role in structuring spatial experience: He claimed that space, as a formal principle of human intuition, "concerns the laws of sensibility of the subject [rather] than conditions of the objects themselves" (Kant, 1929/1770, Sections 15-16).

The same metalogical criticisms apply here as we saw in connection with both Kant's allegedly meaningful notion that the mind actively imposes a temporal order on its experience, and his putatively meaningful conception of things-in-themselves. Here, the purportedly meaningful distinction between human experience, as spatially structured, and "objects themselves," to which the concept of space does not apply, cannot, in principle, be drawn because the conditions such reference requires cannot, in principle, be satisfied.<sup>263</sup>

Were it to be the case, as Kant claimed, that the mind actively imposes spatial structure upon its experience, there would, in principle, be no way in which the preconditions for such reference by the mind to a process of active imposition could be satisfied: What comes to the mind as spatially pre-structured, as condition of spatial experience, does not in principle permit reference to a purported mental process that imposes that spatial structure; Kant's notion is metalogically self-undermining. Once again we confront a delusion-inducing conceptual sleight-of-hand. We shall call projections that lead to this delusion of reference '*projections of spatial constitutive subjective activity*'.

---

<sup>262</sup> "[D]iese Forderung der allgemeinen Kovarianz, welche dem Raum and der Zeit den letzten Rest physikalischer Gegenständlichkeit nehmen...."

<sup>263</sup> This point was previously made in Bartlett (1988).

## 22.5 Projections of temporal and spatial continuity

We turn now to consider a variety of projection encountered in connection with both temporal succession and spatial extension.

We have observed that identifying reference has a temporal basis, and that its temporal basis is logically fundamental to the capacity of a reference frame to identify spatial objects. We've noted that spatial objects are given in perspectival form, generally in a temporally discontinuous manner, and that such objects are characterized by Leibniz boundaries which potentially may, when circumstances are right, be extended, as, for example, when we come to see other sides of objects, sides not initially perceived.

As a result of our perspective-based perception of physical objects, we come to form habitual anticipatory expectations expressed by our common-sense beliefs in the general continuity and regularity of sensory experience. From the standpoint of the metalogic of reference, this simple inferred or imagined continuity, as we've also noted, is unobjectionable. However, as we saw earlier in this chapter in connection with projections of the external world and of other minds, metalogically self-undermining beliefs come about when the inferred or imagined continuity of physical objects is believed to possess an autonomy from frames of reference relative to which such objects can be identified.

Unfortunately, projective beliefs do not end here. It is also a commonly held belief that the serial structure of temporal succession and of spatial order is "everywhere dense"—the notion that between any two elements of a set there exists at least another element. This belief takes two forms, one of which is metalogically self-undermining. On the one hand, there is the belief that (i) temporal succession or spatial order has Leibniz boundaries: This is a belief in the in-principle continuability of a temporal series and the in-principle extendability of spatial perspectives. The anticipatory belief in such continuability or extendability is a belief we commonly come to have: The focus here is this *belief*, one which comes about naturally and habitually, whether by generalization, by the application of principles formulated by a theory, or inductively.

On the other hand, a very different belief can arise; it claims that (ii) the continuability of a temporal series or the extendability of spatial observation is a putatively meaningful expression or manifestation of "an independently existing temporal or spatial continuum." The latter belief can be characterized as an allegedly meaningful belief that temporal succession and/or spatial order are everywhere dense in the metalogically self-undermining sense that between any two elements of a temporal or spatial series of identifiable elements, there always exists, autonomous of what is identified, at least another

element.

De-projected, temporal and spatial continuity *mean* the expectation that forms the basis for belief (i)—whether as a result of habituation or by the application of rules.

The following must be emphasized: Belief (i) is an acknowledgment of the role of a *belief* and is a statement about that belief; it is no more than this. On the other hand, belief (ii) is a claim, not about a belief, but about the putatively meaningful independent existence of serial elements. By the terms of such a claim, such a claim attempts to exceed the referential capacity, in principle, of any reference frame. To claim, not merely that a temporal series or a spatial order possesses Leibniz boundaries, but that such a series or order is “everywhere dense” in sense (ii), is to assert a metalogically self-undermining projection.

This projective conception of the “everywhere dense” nature of temporal succession or of spatial structure is what is ordinarily expressed by the commonsensical, purportedly meaningful notion of temporal or spatial continuity. Such a conception forms, once again, an example of reference gone awry, an instance of a type of claim that undermines its own capacity to refer. We shall call such projections ‘*projections of temporal or spatial continuity*’.<sup>264</sup>

## 22.6 Projections of space-time

In {4.10}, in connection with the self-enclosure of a reflexive, maximally general theory such as the metalogic of reference, I referred, by way of illustration, to the self-enclosed topology of the space-time manifold of a relativistically recurved physical universe. Such an example illustrates how self-enclosure is characterized by an absence of boundaries as is found in a system that is nonetheless closed. In {4.10} I mentioned philosophical systems that may be closed in this sense, forming “sets that are closed upon themselves,” also noting that, for phenomenologists, experience forms a self-enclosed “field” in the sense that no matter how experience is extended, it is never possible to “go beyond” its boundaries.

In the course of this chapter, we’ve seen how Newtonian classical physics postulated allegedly meaningful absolute continua of time and space. On this basis, Newton layered further assumptions—that these absolute continua exist independently of each other, that time passes at an unvaried rate independent of the state of motion of observers, and that spatial coordinates of objects are

---

<sup>264</sup> This projection was first noted in Bartlett (1970, p. 175n). The nature of this projection is related to another and closely allied projection, “projection of the implicit” (see Bartlett, 1974 (Polish) or 1975a).

independently specifiable and autonomous of their time coordinates.

The elimination of the Newtonian notions of absolute space and time by means of the method of de-projection brings with it a recognition that reference has a temporal basis, that the identifiability of spatial objects is temporally based, and that spatial objects are essentially relative to reference frames that permit a retentive record of their serially given past states.

In short, by employing the epistemologically abstract approach of the metalogic of reference, we find a close parallel in the way in which the concepts of time and three-dimensional space are intrinsically interrelated or “fused” in the relativistic conception of physical space-time; we come to see that spatial reference is necessarily related to its temporal basis. In this sense, it is appropriate to understand identifying reference to physical objects in essentially conjoined spatio-temporal terms.<sup>265</sup>

If we do this, habitual projective ways of thinking may lead to the purportedly meaningful belief that, corresponding to a spatio-temporal framework of reference that provides a basis for reference to physical objects, there purportedly exists “an independently existing physical space-time continuum” in which those objects are “located.” Here, once again, we encounter a belief that seeks to trespass beyond its metalogical horizon. It is only in essential relation to sets of physical objects, which are given perspectively in temporally based reference frames, that their spatio-temporal identifiability is possible. We shall call projective beliefs of this kind ‘*projections of space-time*’.

## 22.7 In retrospect

This chapter has identified a series of self-undermining projections related to the concepts of time, space, and space-time. I have not sought to give an exhaustive enumeration of such projections; rather, my intent has been to describe representative, prototypical projections that belong to this group. Let us consider in retrospect the projections that have been discussed. They are:

Projections of time:

- (1) *Projections of the past*
- (2) *Projections of time-flow*
- (3) *Projections of the future*
- (4) *Projections of absolute time*

---

<sup>265</sup> This conclusion is reached from a phenomenological standpoint in Bartlett (1970, Section 2.1, “Phenomenological Space-time”).

(5) *Projections of temporal constitutive subjective activity*

Projections of space:

(6) *Projections of absolute space*(7) *Projections of spatial constitutive subjective activity*

(8) Projections of temporal and spatial continuity

(9) Projections of space-time

Certain of these nine varieties of projection are especially resistant to attempts to eliminate them from one's everyday conceptual vocabulary. As noted before, to incorporate the results of de-projective analysis in one's thinking can often require hard work to offset their often counterintuitive nature.

Much philosophical analysis is technically specific in its application to highly delimited topics, rendering it of little consequence to an individual's general outlook, which traditionally was called the individual's '*worldview*'. The analysis provided in this chapter may certainly be regarded in this technically restricted way. But to do so would be to mistake the author's purpose and to miss the opportunity of reaching what is likely to be a new and intellectually challenging transformation of conceptual outlook on the part of the reader.

Unfortunately it is not possible or to be expected that the written lines of a book can, of themselves, bring about such a fundamental change in even the most attentive, sympathetic, and intellectually competent reader. Change of this kind and magnitude tend only to come about as a result of the reader's own efforts to integrate and apply within his or her own thinking the results that have been reached.

Notwithstanding these unavoidable limitations, as this study's author I would like to sketch—albeit, in the brief space available, somewhat in caricature and impressionistically—how one might bring together the de-projective temporal and spatial results we have reached.

The elimination of the first three varieties of projection (of the past, of time-flow, and of the future) results in an understanding of the essential relativity of temporal events to the observer's frame of reference, and usually this means the essential relativity of temporal events to the observer's own experience. The elimination of these projections brings with it a recognition that past events, the perception of change of time, and future events are, each of these, a function, respectively, of memory or other means of record-keeping, of the temporally successive nature of perspectively given information, and of the observer's anticipation or expectation of what is likely to occur—whether these are formed by habituation, or by virtue of a rule or law in which

confidence has been invested, or by probability estimations and predictions.<sup>266</sup> The challenge, given the deeply rooted nature of these projections in commonsense thinking, is to recognize that the possible meaning of the reality of “the past,” “the passage of time,” and “the future” is no more than what has just been said; these notions do not refer to “more than this.”

The further de-projective elimination of the projections of absolute space and of absolute time lead to a recognition that “space” and “time” are—and are no more than—forms of serial order given in the observer’s experience. These concepts have no possible meaning independently of the reference frames in terms of which temporal and spatial identifying reference is possible.

As we then continue by eradicating the projections of temporal constitutive subjective activity and of spatial constitutive subjective activity, we avoid the self-undermining putatively meaningful Kantian notions that the observer’s mind or subjectivity “imposes” the temporal and spatial ordering structure experienced by the observer. We instead recognize that temporal and spatial ordering is ingredient or constitutive of possible reality for the observer. We do not make the additional, projective, step that claims that such reality is somehow purportedly “due to subjective activity or processes.”

Once the further projections of temporal and spatial continuity are dismissed, we understand that continuity of either sort is a function of what is given perspectively; continuity that “fills in the blanks” of periods of discontinuous observation is a continuity that is conceived or imagined; it cannot meaningfully be claimed to have possible autonomous reality beyond this.

We then come to projections of space-time. When such projections are de-projectively eliminated, we see that spatial reference has a necessary temporal basis: Identifying reference to physical objects is fundamentally spatio-temporal; we recognize that the order structure of such objects is serial in nature, where by ‘serial’ we understand successive, reiterable reference that makes it possible to refer to “the same” object. The de-projective concepts of time and space have, then, a common basis in the principle of serial order: Spatial order embodies the serial nature of temporal order. The identifiability of physical objects is temporally based; we recognize that physical objects are essentially relative to reference frames that provide a means of record-keeping of their serially given past states. The putatively meaningful notion of “an independently existing physical space-time continuum” is rejected as meta-

---

<sup>266</sup> In this sense, change is more fundamental than time, and as a result “time” is a construct that simply measures change. We know what change is, but ask, What is time? If this question asks for something that “exists independently of change,” it is a projectively meaningless question. We know what distance is, but then ask, What is “space”? It is the same there.

logically self-undermining. Instead, we see that the identifiability of physical objects and of the time of occurrence of events is framework-relative: A reference frame that permits such identifying reference presupposes spatio-temporal serial order, which has no applicability independently of such a frame of reference.

The view that results from the de-projective elimination of all nine varieties of projection described in this chapter is a view that is wholly “intrinsic” in the sense that temporal and spatial order are to be understood as no more than forms of serial order to which reference is made possible when appropriate reference frames are employed. “Time,” “space,” and “space-time” have no possible meaning except in this intrinsic framework-relative sense.

The conceptual challenge in integrating these results—that “time,” “space,” and “space-time” are “no more than this,” that they are not, in some inchoate way, “more” than forms of order that are simply and only functions of the reference frames we employ—is difficult thanks to the pervasive and deeply rooted human projective disposition which, to use Kant’s word, “compels” us, or to use Graham Priest’s, “drives” us to trespass beyond horizon boundaries. It is, as we’ve now seen in a wide variety of contexts, an urge to think, to believe, and to claim that “there is more beyond,” that “there is a ‘Far Side’” that lies beyond possible capacities of knowledge—not only beyond human capacities, but beyond those of any knower. The commonsense, intuitively persuasive notions of “time,” “space,” and “space-time” are among the most widespread, habitual expressions of this compulsion to engage in self-undermining projections.



## 23

### **The Projections of Causality, Determinism, and Free Will**

Already in this study we have examined a number of conceptually “short-circuiting” assertions that involve relational ascriptions, i.e., claims that utilize concepts which involve such relations as “is independent of,” “is dependent upon,” “is influenced by,” or “is determined by.” We have found that projections result when a wide variety of putatively meaningful claims about these relations are made, as in the case of projections of discovery or invention in problem-solving, mathematics, and physics; in connection with projections relating to the conceptually unreachable; in the case of the notions of an external world, things-in-themselves, other minds, realism, and idealism; and, in the last chapter, in connection with the various forms of temporal, spatial, and spatio-temporal projections.

In this chapter, we shall look at still another group of projective concepts and claims that involve a different variety of relational ascription, one that asserts either that there is a causal relationship that connects two or more events or phenomena, or that such a relationship does not exist so that the events or phenomena are thought to comprise expressions of “free will.”

#### **23.1 Causality**

In keeping with the objectives of a “negative science” ({12, 13.5, and 15.2}), it will not be my intent in this chapter to provide a unitary definition of the cause-and-effect relationship as it is exemplified in its many forms. Others have tried to do this, with different degrees of success, a project that requires at least a book in itself. Instead, the two-fold objective here will be to identify a number of ways in which the commonly accepted notion of causality is metalogically self-undermining, i.e., is projective in the metalogical sense with which we have become familiar, and then to provide a de-projective revisionary re-formulation of that concept.

There are many distinguishable conceptions of causality. These include Aristotle’s four notions of efficient, material, formal, and final causes, as well

as later notions of a first cause, secondary causes, necessary causes, sufficient causes, necessary and sufficient causes, contributory causes, counteracting causes, countervailing causes, legal causes, proximate causes, predictive causes and probabilistic and Granger causes, and even “negative causes” that are thought to function as a result of their absence. To develop a definition of the cause-and-effect relation capable of embracing such a many-faced menagerie would likely be simplistic and uninformative. Nonetheless, for our purposes we shall need a basic notion of causality with which to begin—a notion of the relation between cause and effect that expresses the intuitive, common-sense conception which, over the centuries, has become conventional, habitual, and seldom questioned. The notion we shall sketch is one that is embraced daily by most people, is used uncritically by many specialists in their respective disciplines, and yet is a notion which we shall find is conceptually primitive and rationally unacceptable. We shall refer to it as the ‘conventional notion of causality’.

### 1. *The conventional notion of causality*

In its widely used conventional sense, causality is considered to be the connection that is believed to exist between a cause, *C*, which may be an event, a phenomenon, a process, or other identifiable object of reference, and *E*, a distinguishable second event, phenomenon, process, or other identifiable object of reference. The conventional notion of causality makes the following group of ordinarily imprecise and open-ended claims about the connection between a cause, *C*, and an event, *E*:

- (1) *C* usually occurs earlier in time than *E*;
- (2) *C* is responsible for bringing *E* about;
- (3) the occurrence of *E* is in some way dependent upon *C*;
- (4) *C* and *E* are regularly and uniformly associated with one another in that whenever *C* occurs, *E* occurs;
- (5) etc.

Condition (5), the “etc.” in this group, plays an important role since conditions (1) – (4) are each subject to qualifications and stipulations such as “all other things being equal,” “there are no other intervening or interfering conditions,” etc. —Here, too, a listing of qualifications and stipulations like these must be followed by “etc.” These situationally relevant “etceteras” will not concern us; their exhaustive enumeration will not be important here, and we may omit their discussion.

In addition to the above conditions, the conventional notion of causality brings with it what we might think of as a second layer of claims, which it will be convenient to distinguish as *beliefs about the relation between cause and effect*. They include:

- (6) *C* and *E* are believed to be, in certain significant respects, “similar” to one another: For example, both are often believed to belong to the same category of events (or phenomena, processes, or other identifiable objects of reference), or they are believed to possess certain defining properties or characteristics in common, etc.
- (7) The causal relationship between *C* and *E*, described using words like ‘responsible’ (condition (2)) or ‘dependent’ (condition (3)), is believed to be a “necessary” or “indispensable” relationship. How this conventional notion of “necessity” or “indispensability” is to be understood will concern us in some detail.
- (8) The “regularity” and “uniformity” mentioned in condition (4) are believed to be sufficiently persuasive so as to supply, or to add, sufficient justification to support confidence in belief (7).

## 2. *The experiential basis of the conventional notion of causality*

It has been hypothesized that the conventional notion of causality as we have so far described it has its origin in the experience of effort-and-resistance when a person tries, for example, to move physical objects—in other words, the feeling of muscular effort.<sup>267</sup> Alternatively, the conventional notion of causality may originate in certain early childhood experiences; perhaps it is wholly or partially learned; or it may be a combination of these. Whatever its *developmental* origin, its *basis* as a set of beliefs—the center around which causal beliefs are expressed and acquire their convincingness—is most plausibly experiential: For example, we *feel* a physical relationship between pushing a cart, the resistance it gives, and its consequent movement or

---

<sup>267</sup> Physicist Max Jammer (1957, p. 124), in speaking of the notion of force, commented that force is “a concept given a priori, intuitively and ultimately in analogy to human muscular force.” Similarly, sociologist John Levi Martin claimed: “It is ... when our body acts on external objects that we are most likely to sense causality.... I will refer to this understanding of causality as the root of the ‘commonsensical’ one, as it seems that this interpretation is indeed common in everyday life” (Martin, 2011, p. 32; for references to other authors who have held this view, p. 31).

stubborn immobility. We *hit* a nail squarely on its head and see the nail penetrate more deeply into the wood. Here, *C* and *E* are *physically experienced*; they are perceived to be temporally and spatially contiguous; we recall a regularity and uniformity of their sequential association; and we come to hold a set of beliefs about the relationship between them. The experiential basis is what it is, and as a set of facts offers little for us here to be concerned about. However, it is a very different matter when it comes to the “second layer” of beliefs that we typically affirm about that experiential basis, beliefs which are responsible for philosophical problems that have arisen in connection with the conventional notion of causality.

### 3. *The conventional notion of causality considered philosophically*

It will be helpful to distinguish the following topics related to the conventional notion of causality, each of which, for centuries, has served as a source of very active philosophical discussion and contention:

- ◆ the temporal nature of causal relations
- ◆ the recognition of similarity relations between cause and effect
- ◆ the belief that the relation between cause and effect is one of “necessity” or “indispensability”
- ◆ the associated belief that the relation between cause and effect involves some kind of “tie,” “agency,” or the transmission of “productive power”
- ◆ the belief that the past regularity and uniformity of a cause-and-effect relation can be relied upon in the future

Let us consider these one at a time:

### 4. *The temporal nature of causal relations*

The conventional notion of causality generally assumes that causes occur before their effects (as claim (1) above asserts). But this is not always the case: For example, the commonsense view of causality accepts the view that the Sun’s gravitational field “causes” the planets to maintain their various orbits, despite the fact that here the “cause” and the “effect” are not temporally successive. In fact, as we shall note later, many invariant laws of nature are described in functional terms which have no need to include reference to time.

Notwithstanding the fact that, from the standpoint of the conventional notion of causality, not all causes are required temporally to precede their effects, the conventional notion presupposes a framework of reference in terms of which it is possible identifyingly to refer to a cause as distinguishable from its effect. Claim (1) meets this condition since, in most commonly encountered situations, cause and effect are temporally successive and are distinguishable at least because of this. But when a functional relationship is in view, as in the case of the gravitation example above, the presupposed reference frame provides a basis for distinguishing the condition of gravity from its functionally defined effect, the maintenance of the orbits of the solar system's planets, but here the factor of temporal precedence of the cause in relation to its effect does not play a role.

Conventionally viewed causal relations must then presuppose one or more reference frames in terms of which it is possible, in contextually relevant ways, to discriminate causes and their associated effects. For this to be possible, we need to recall that reference itself has a temporal basis.<sup>268</sup> Reference to any object of reference is only possible in terms of some presupposed framework permitting that object's re-identifiability. Alternatively expressed, such a framework must make it possible to retain a record of temporally earlier states; such a framework must provide for the capacity for sequential or serial retention of earlier states. In terms of the conventional notion of causality, no matter whether a cause and an effect are temporally successive or whether their relationship is functionally defined, identifying reference to them metalogically presupposes a reference frame's capacity to retain and to refer to temporally earlier states. Observations of the regularity and uniformity of causal relations satisfy this requirement, for without the capacity to retain information about past associations of causes and effects it would of course not be possible to take note of their regularity and uniformity.

##### 5. *Similarity relations between cause and effect*

[W]e may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.*

– David Hume (2004/1748, Sec. VII, original italics)

This was one of two definitions Hume gave of the relationship between cause and effect. This definition pointed to the conceptual fact, not analyzed in

---

<sup>268</sup> See {10.1} and {22.1}.

detail by him, that, in order for us to take note of a pattern of regularity and uniformity among causes and their effects, we must be able to recognize similarities between objects. That is, we must be capable of recognizing that the causes that we observe are of certain kinds, and that the effects are of certain kinds, and then be able to refer to what we recognize these to have in common. Otherwise, we should not be able, in principle, to take note of patterns of regularity of similarly occurring causes and subsequent effects (Hume assumed that causes occur before their effects).

This has become one of the main conceptually fundamental ways of understanding the cause-effect relationship: Similar causes can be grouped together because they bear certain relationships of similarity to one another, as can similar effects; causes and their corresponding effects also bear certain similarity relationships to one another which allow them to be recognized as linked; and, as a result of these relationships, their paired association can be distinguished from the entirety of other phenomena or events.

This recourse to and reliance upon the notion of similarity, however, encounters serious and largely unrecognized conceptual deficiencies whose study lies outside the focus of this work. Nonetheless, since these deficiencies add to the justification provided in this chapter that shows that we must relinquish the commonsense notion of causality, a note purely to acknowledge those conceptual shortcomings is made here. They relate to the inescapable ambiguity of any attempt to claim that identified similarities are uniquely determined. The role of *selection* in identifying similar causes in relation to similar effects must disengage these from a general background of the great multiplicity of other possible phenomena or events, and in doing this, is unavoidably burdened by necessary inescapable ambiguity.<sup>269</sup>

#### 6. *The belief that the relation between cause and effect is one of “necessity” or “indispensability”*

[W]e may define a cause to be *an object followed by another*

---

<sup>269</sup> For a detailed analysis, see the author’s monograph, Bartlett (2015).

This problem, as it relates to an analysis of the commonsense notion of causality, was tangentially touched upon but then simply set aside by Lewis (1973, pp. 558-559) when he wrote: “We sometimes single out one among all the causes of some event and call it ‘the’ cause, as if there were no others. Or we single out a few as the ‘causes,’ calling the rest mere ‘causal factors’ or ‘causal conditions.’ Or we speak of the ‘decisive’ or ‘real’ or ‘principal’ cause. We may select the abnormal or extraordinary causes, or those under human control, or those we deem good or bad, or just those we want to talk about. I have nothing to say about these principles of invidious discrimination.”

... *where, if the first object had not been, the second never had existed.*

– Hume (2004/1748, Sec. VII, original italics)

This was Hume's second definition of causality.<sup>270</sup> It was an attempt to explain the conventional belief that the relation between cause and effect is some indistinctly understood variety of "necessity" or "indispensability." In the nearly three centuries since then, an endless stream of philosophical publications has continued to pour forth, dissecting the nature of this putatively meaningful, "contrary-to-fact" (or "counterfactual") conditional relationship between cause and effect. The conventional notion of causality is deeply invested in the belief that the cause-and-effect relation is one of *indispensability*, of some kind of *necessity*—that, if the cause had not happened, the effect would not either.

In many instances, of course, the effect does happen, but without the cause that can lead to it: Cardiac arrest can cause death, but death often occurs without that cause. In narrowing down the factors recognized as relevant, we eliminate or disregard those occurrences judged to be irrelevant, and focus specifically on the situation in which a certain cause happens and a certain effect occurs regularly and uniformly and can be correctly predicted. Once we believe we have successfully accomplished this, the relation between the cause and the effect is most commonly considered to be a counterfactual relation: The specified effect would not, it is conventionally believed, occur had not the effect happened.

Is such a belief possibly meaningful? Can the referential conditions necessary in order for it to be possible to refer to what "would have happened had not what actually happened, happen?" What are the referential preconditions that must be satisfied in order for reference to be possible under such putatively meaningful circumstances? Do such circumstances rule out the very possibility the satisfaction of that which they require?

Certainly we can, and do, refer to a description of circumstances that did not, or do not, actually come about. When we do this, we do not make an effort to stretch beyond the metalogical horizons of our frameworks of reference. Such descriptions are offered as alternatives to what has happened at a past time, or has generally happened in the past. References to such imagined

---

<sup>270</sup> He ran the two definitions together: "[W]e may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.* Or, in other words, *where, if the first object had not been, the second never had existed.*"

alternatives are straightforward and do not give rise to metalogical difficulties.

But if, in connection with what we take to be a cause-and-effect relation between *C* and *E* at time *t*, we make the claim, *P*, that “*E* would not have happened at time *t* if *C* had not happened,” we catch ourselves attempting to refer in a way that abrogates, nullifies, or, as we’ve often expressed this, undermines the very possibility of such reference. By including the time specification, *t*, satisfying the referential preconditions for reference to *P* is very evidently ruled out—unless we entertain alternative possible worlds, which, of course, we are free to do. However, attempting this would miss the point: If *P* putatively refers to *this* world, not some alternative universe, the referential preconditions of *P*’s purported meaningfulness “short-circuit” or “implode”: It is—necessarily—nonsensical to advance a claim like *P*. In principle, the conditions that must be satisfied for *P* possibly to refer cannot be satisfied.

This is the first variety of causal projection we have encountered in this chapter; we shall call projections of this sort ‘*projections of counterfactual causation*’.

As we shall see a little later, rigorous science has come to eschew such projective claims, but they do persist, often very vigorously, in causal assertions, frequently made with great confidence, both by advocates of common sense, by many philosophers, and by many professionals in other disciplines. In behavioral science, for example, the counterfactual notion of causality is now widespread. Indeed, not very long ago, John Levi Martin (2011, p. 33) commented “a counterfactual definition of causality is now dominant across the social sciences....”

7. *The relation between cause and effect involves some kind of “tie,” “agency,” or the transmission of “productive power”*

Let us continue our examination of the conventional notion of causality. It is characterized by the commonsense belief that a vital “*tie*” exists between cause and effect. It is a belief that between cause and effect there exists some very close, quasi-mechanical, energetic push-pull connection, reminiscent of the feeling of muscular effort, a connection that is established between the two so that the one exercises a *productive* or *generative power* or *force* that brings about changes in the other, or actually brings the effect into existence. This is the putatively meaningful belief that some sort of “force” is transmitted from cause to effect that brings the effect about.

It was Hume’s incisive, no-nonsense observation that “... *we never can, by our utmost scrutiny, discover any thing but one event following another... One event follows another; but we never can observe any tie between them*” (Hume, 2004/1748, Sect. VII, Pt. II, italics added).



According to Hume, not only are we *incapable* of observing any such “tie,”

We *deceive* ourselves in imagining we can form any such general idea [of the “power” or “necessary connexion” between cause and effect]... [W]e have really no distinct meaning, and make use only of common words, without any clear and determinate ideas.... [I]f we go any further, and ascribe a power or necessary connexion to these objects, this is what we *can never* observe in them.... (Hume, 1888, I, III, 14, italics added)

[W]e have no idea of connexion or power at all ... *these words are absolutely without any meaning*.... (Hume, 2004/1748, Sect. VII, Pt. II, italics added).

Hume does not explain precisely and convincingly why we “*never can*” (or “*can never*”) discover any more than one event following another. I’m sure he felt this was an obvious and convincing fact about experience, but it is a good deal more than this.

Let us ask ourselves what referential conditions would need to be satisfied in order for us to be able, in principle, to detect such a connection of “force” or “agency” or “generative power” either in a cause itself, or in its relationship with its associated effect. If it should occur to us to think that such a “force,” “agency,” or “generative power” can be recognized by virtue of a conditional contrary-to-fact relationship between cause and effect (if the cause had not occurred, the effect would not have), then we must dismiss this as an instance of a projection of counterfactual causation, and hence devoid of possible meaning.

Alternatively, we might look for information that tells us something about what may lie intermediate between cause and effect. Suppose we found such information; it would comprise an object of reference—an intermediate event, phenomenon, or datum to which we can refer. But then we face the same, and endlessly recurring, question: What referential conditions must be satisfied in order to detect that such an intermediate event comprises a “force,” “agency,” or “generative power” that is “transmitted” to the next link in the causal chain, reaching its culmination in the effect?

As Hume recognized, “... we never can, by our utmost scrutiny, discover any thing but one event following another....” This is not because we lack a skill that would permit such a discovery; the very conditions such a possible

discovery would require cannot, in principle, be met.

Attempts to claim more than this exceed the referential capacity of any reference frame. They are instances of a second variety of causal projection. Projections of this kind we shall call '*projections of genetic causation*', alternatively, '*projections of causal agency*'.

In addition to a persuasive record of consistent regularity and uniformity of a cause-and-effect relation—however law-like their correlation and predictability may have become—established correlations between cause and effect *can, in principle, provide no possible referential ground* for the view that a cause “necessitates,” “compels,” “forces,” “produces,” or “generates” its associated effect.<sup>271</sup>

To my knowledge, the philosopher of the past has come closest to this recognition was Moritz Schlick when he wrote:

[Philosophers] ... are perfectly wrong when we think that [a causal] chain could consist of anything but events, that it could be a kind of mysterious tie called “causality.” The conception of such a “tie,” which is really not a concept but a mere word, *is due to a faulty process of thinking* that is very common in the history of philosophy: *the continuation of a thought beyond its logical limit; we transcend the region in which a word makes sense and necessarily find ourselves in the region of nonsense.* (Schlick, 1949/1931, p. 522, italics added)

Schlick reached and asserted this conclusion presuming the logical positivist's bias in favor of verificationism; he claimed this conclusion was necessary, but did not demonstrate it. Like Hume, he recognized that the notion of causal “power” or “necessary connexion” leads to meaninglessness. Had he not been murdered by a student, he might have developed a proof that there are indeed “logical limits” to thought, limits which, if transcended, lead us into a “region of nonsense.”

8. *The belief that the past regularity and uniformity of a cause-and-effect relation can be relied upon in the future*

The connection of experienced past uniformity with expectation as to the future is just one of those uniformities of

---

<sup>271</sup> This observation was first made in Bartlett (1970, Sect. 2.5).

sequence which we have observed to be true hitherto. This forms a psychological account of what may be called the animal belief in causation, because it is something which can be observed in horses and dogs, and is rather a habit of acting than real belief. So far, we have merely repeated Hume, who carried the discussion of cause up to this point, but did not, apparently, perceive how much remained to be said.

– Bertrand Russell (1972/1914, p. 220)

When a set of events (phenomena or other objects of reference) has come to be understood in terms of law-like rules that make it possible to predict the occurrence of a second set of events, the conventional notion of causality considers that the law-like regularities that have been identified between the two sets comprise causal relations that *explain* the occurrence of events in the second set. As we have seen, this “causal explanation” is conventionally believed to involve similarity relations between cause and effect, a connection of “necessity” or “indispensability” between them, and a “tie,” “agency,” or “productive power” that is transmitted from cause to effect.

Once these conventional beliefs are relinquished because of their meta-logically self-undermining nature, we are left with a schema of representation which is based upon informal sets of expectations among common folk or upon formal theories among scientists, which takes into account the past record of regularity and uniformity of a cause-and-effect relationship, and, on the basis of this record, which provides what is considered to be a convincing ground for reliable predictions in the future. We remain always prepared to revise our predictive expectations in the light of future evidence, but despite this caveat, we gain confidence in the reliability of our expectations in proportion to the degree to which they have been confirmed in the past. Whether these are mere intuitive expectations or predictive theories to which they give rise, they are inherently descriptive generalizations formed on the basis of past experience. In relinquishing the projective conventional beliefs in counterfactual and genetic causation, we implicitly shift from an *explanatory* framework to one that is purely *descriptive*, as we shall make clear.

The question whether the confidence we invest in our causal expectations is justified leads directly to the topic of inductive reasoning and to the philosophical problem whether it is justifiable to assume that the consistency of past observed regularities is predictive of the continuation of those regularities. If the latter is justifiable, then what has come to be known as the princi-

ple of induction is defensible. Whether this can be done is one way of stating the philosophical problem of induction.

As Russell concluded at this juncture:

It is thus the principle of induction, rather than the law of causality, which is at the bottom of all inferences as to the existence of things not immediately given.... Whether inferences from past to future are valid depends wholly, if our discussion has been sound, upon the inductive principle: if it is true, such inferences are valid, and if it is false, they are invalid. (Russell, 1972/1914, p. 226)

The problem of induction raises a pair of questions: whether it is justifiable to extend generalizations made on the basis of a group of past observations beyond what has been observed, and, more specifically, whether it is justifiable to believe that cause-and-effect relations regularly and uniformly observed in the past will continue to exhibit the same regularity and uniformity in the future.

The framework-relative methodology which this study has shown to be strongly compelling immediately leads to the following result: Both of the above inductive questions purport to raise meaningful questions whether, in most general terms, we are justified either in believing or in claiming with certainty that Leibniz boundaries of certain kinds can reliably be counted upon to be extendable in the future as they have in the past, i.e., in a manner that consistently maintains past observed correlations of causes and their associated effects.

There are two possible responses to these questions, one relating to the justifiability of a *belief*, and the other to the justifiability of a claim to *certainty*. The belief that, as we walk around a three-dimensional physical object, we will perceive sides or aspects of the object previously not observed is a belief most often repetitively confirmed by our sensory experience (we of course recognize it is only a belief since sometimes we can be misled by two-dimensional cut-outs that appear to be three-dimensional, or by other illusions). A repetitively confirmed belief of this kind comes habitually to be relied upon. Here, it is reasonable to speak of the role of *conceptual and psychological habit*. This was the conclusion reached by Hume: It is by virtue of “custom” or “habit” that we come to expect past cause-and-effect uniformly observed regularities to persist in the future: “without the influence of custom we would be entirely ignorant of every matter of fact beyond what is immediately present to the memory and senses” (Hume, 2004/1748, Sect. V, Pt. II).

However, when it comes to the presumably meaningful question concerning the justifiability of a corresponding claim to *certainty*, we not only find our hands empty, we find ourselves in a conceptually delusional state, seeking to grasp what escapes any possible grasp: There is no referential basis, in principle, in terms of which this putatively meaningful question *could* make sense. Understood in this specific way, the problem of induction as it applies to causality is inherently projective. We shall call projections of this sort '*projections of framework-transcending causality*'.<sup>272</sup>

In short, the alleged "problem of induction" as it applies to causality ceases to be a meaningful philosophical problem. On the one hand, we accept the legitimacy and the usefulness of habit and of the expectations, in the light of past experience, in which habit leads us to invest our confidence. On the other hand, we recognize that if carried beyond this, the "problem of induction" leads directly to projections of framework-transcending causality.

Karl Popper, for example, reasoned that the problem of induction asks the wrong, or an improper, question: It asks, he thought, how to justify theories that are such that they cannot be justified inductively. Popper wrote, "I approached the problem of induction through Hume. Hume, I felt, was perfectly right in pointing out that induction cannot be logically justified" (Popper, 1963, p. 55).

But Popper's reasoning misses the point: It is not merely that induction cannot be logically justified, it is rather that the putatively meaningful question whether it can be justified to consist in more than general expectations formed from past observational patterns, is to ask a metalogically self-undermining question. It is, in yet another guise, the attempt to reach beyond the reachable, to transgress the metalogical horizons of possible frameworks of reference.

### 23.2 The evolution of the concept of causality

The law of causality ... is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm.

---

<sup>272</sup> A second example of this projection is found in the purportedly meaningful question "whether there is a cause 'outside' the physical universe." It is reasonable to call such a projection 'extra-systemic', 'extra-experiential', or 'extra-phenomenal', but it will be appropriate and convenient simply to include it under the above name, 'projections of framework-transcending causality'.

[I]n a sufficiently advanced science, the word ‘cause’ will not occur in any statement of invariable laws.

– Bertrand Russell (1912-13, p. 1; 1972/1914, p. 223)

Early in this chapter, I referred to the conventional notion of causality as “conceptually primitive” and “rationally unacceptable.” So far in this chapter we have taken note of three varieties of projection that are commonly involved in the conventional notion of causality: projections of counterfactual causation, projections of genetic causation (or of causal agency), and projections of framework-transcending causality. To these three varieties of causal projection we shall shortly add a fourth, “*projections of hidden determinants*,” which, as we shall see in a later chapter, has played an important role in the evolution of quantum theory. To the extent that many everyday people as well as many specialists embrace and make use of the conventional notion of causality that involves these metalogically self-undermining projections, to that extent we are justified in judging the ordinary, commonsense notion of causality to be conceptually primitive and objectionable to reason. It is also primitive from the standpoint of historical development.

Historically, the notion of causality has evolved in three stages: In both the commonsense view and in classical physics, causes have traditionally been thought generally to precede their effects. With the development of Einstein’s special theory of relativity, the idea of “causal precedence in time” was revised to mean that causes must occur in the past light cone of any effect, and any effect must occur in the future light cone of the cause. But with Einstein’s development of the general theory of relativity, the concept of cause-and-effect relations was dissolved and its place has been taken by principles of invariance, i.e., physical laws that are formulated in terms of *general functional dependency*, which we consider in detail in the next section.<sup>273</sup> When expressed in terms of the invariance of functional relationships among events, we find that physical laws frequently do not require reference to time variables, and as a result the traditional condition of the temporal priority of causes in relation to their effects is gradually also being left behind. With the advent of quantum theory, as we shall see in {27}, the conventional notion of causality has similarly been relinquished in favor of a *functional* understanding of relationships among events relative to their observability.

By giving up and by rejecting the projective conventional beliefs in counterfactual, genetic, and framework-independent causation, an implicit shift, as

---

<sup>273</sup> Special and generality relativity are treated in {26}.

noted earlier in passing, is made from an *explanatory* framework to one that is purely *descriptive*. We shall now be able to make this clear.

### 23.3 The functional understanding of causality

Causality means nothing but a functional dependency of a certain sort. We must emphasize this because time and again the opinion is advanced that aside from the functional dependency between two events, there must be a “real” relation or “essential relation,” namely, such that the first event “produces,” “generates” or “brings about,” the second. It is strange that the opinion is still held even by physicists and epistemologists that science ... must not rest content with an investigation of those functional dependencies, but that it should ascertain, above all, the “real causes.”

– Rudolf Carnap (1967/1928, p. 264)

Philosophers and psychologists alike have continued to suppose that science must aim primarily at discovering causes, although as a matter of fact the word ‘cause’ vanished long ago from the vocabulary of the more advanced of the sciences....

[W]e ... adopt the position that, although causal links are not discoverable, *implications of probability* are, and these will enable us to predict as well as to describe without assuming any external causal compulsion.

[T]he word ‘causes’ ought in every rigorous argument to be carefully eschewed, and be replaced by some non-committal description such as that of “functional relations.”

– Cyril Burt (1940, pp. 219, 220n, 234)

Not everyone, of course, agrees with these anti-conventional conclusions about causality. A few pages back, I quoted John Levi Martin’s observation that the counterfactual notion of causality today dominates the social sciences. J. L. Mackie similarly commented, and with approval:

[W]hereas Russell thought that causality was out of date in 1912, causal concepts are, sixty years later, constantly being

used in our attempts to understand perception, knowledge, and memory, and to clarify our thought about action, responsibility, legal claims, purpose, and teleology. (Mackie, 1974, p. 154)

What Mackie claimed in the early 1970s remains true today: Despite the evolution of the notion of causality witnessed in physics, the projective, commonsense, conventional notion of causality continues, both in daily life and in many disciplines, to seduce through its conceptual sleight-of-hand, with a singular disregard for assertions that trespass beyond the horizons of their possible meaningfulness. It is time to dispense permanently with the conceptually primitive notion of causality, to recognize its in-principle meaninglessness, and to replace it with a coherent and clear concept. As we shall see, a functional understanding of causality accomplishes this.

The passage quoted from Carnap at the beginning of this section claims that cause-and-effect relations are “nothing but a functional dependency of a certain sort.” Much rides on the meaning of the phrase ‘*functional dependency*’ since it is the relation of “functional dependency” that comes closest to the commonsense notion of causality in the de-projectively revised concept we shall develop. The relation of the functional dependency between cause and effect provides a basis for understanding how an identified cause may be considered “responsible” for its associated effect, and how that effect may be thought to “depend” on the cause. The relation of functional dependency supplies this—but without introducing projective claims about counterfactual, genetic, or framework-transcending causation. Let us see how this is the case:

The clearest and most directly relevant concept of a function is found in mathematics. A mathematical function is an abstract formalization of a relationship that expresses how variations of one kind are associated, or depend upon, variations of another kind. Understood in this general sense, a function expresses a relation that associates each element of a set  $X$ , called the ‘domain’ of the function, with an element in a set  $Y$ , called the function’s ‘codomain’ (which is sometimes the same set as or may have members that belong to  $X$ ). The relation expressed by a function may be simple, involving two distinguishable sets, or it may be complex, involving multiple sets. But whether simple or complex, the relation expressed by a function can be understood to be one that defines the *dependency* of the association of each element of one set upon a corresponding element of the other set. (Similarly, the relation can be understood to define how the association of one element is *responsible* for the other.) Functional dependency, then, describes a relation between elements of distinguishable sets, and it clearly has nothing whatever



to do with one element “producing” another, and nothing to do with a counterfactual claim of “what might happen if one element did not occur.” If  $x$  and  $y$  are elements belonging to different sets, if  $y$  functionally depends upon  $x$ , and if  $x \supset y$ , then it makes sense to say, though somewhat impressionistically, that  $x$  “is responsible for”  $y$ .

Historically, physics has provided the clearest, most readily employed and built upon understanding of cause-and-effect relations, and so it can be helpful to look again to physics for signs of progress. In physics, cause-and-effect relations are most commonly expressed in the form of physical laws. Physical laws have the following general properties: They are universal statements that summarize functional relations between factors judged to be salient—factors, that is to say, which have been selected in the belief, and supported by past or future expected observational evidence, that they are fundamental and indispensable to the functional relation expressed by a given physical law. Physicists attempt to find the simplest generalized formulation of a functional dependency that is invariant (i.e., is preserved by physical or mathematical transformation operations) and is universal and therefore independent of the particular, contingent conditions of individual reference frames.<sup>274</sup> In addition and fundamentally important, the functional dependency expressed by a physical law must ideally be testable by means of falsifiable predictions made on its basis.

Physical laws are most frequently expressed in the form of equations that formulate the functional *interdependency* of both sides of an equation. A familiar physical law that expresses this is Newton’s second law,  $f = ma$ , which formulates the relation of functional interdependency between force, on the one hand, and, on the other, the mass of an object and its change in velocity (i.e., its acceleration).<sup>275</sup>

Physical functional interdependencies expressed by means of equations make it possible to predict the consequences of varying an equation’s factors. A great deal of weight is placed on those consequences which are observable, consequences which can confirm or falsify the functional interdependencies asserted by an equation. When such predictions are made, they may—in the de-projectively revised sense developed here—be thought of as the “*causal consequences*” of the given variations. Such “causal consequences” are *impli-*

<sup>274</sup> See {4.3}. In relativity physics, for example, the invariance sought must meet the requirement that physical laws are to take the same form for all observers (see {26}).

<sup>275</sup> Other familiar examples of functional interdependencies include Newton’s law of universal gravitation,  $F = G(m_1m_2)/r^2$  (expressing the functional interdependency between, on the one hand, the gravitational force,  $F$ , between two objects, and, on the other, their respective masses,  $m_1$  and  $m_2$ , the distance between the centers of their masses,  $r$ , and Newton’s gravitational constant,  $G$ ), and of course Einstein’s celebrated mass-energy equivalence,  $E = mc^2$ .

*cations* that are explicitly *theory-based*: They should be recognized as essentially *framework-relative* implications; they form the set of the predictable consequences of a theory. In this connection, we recall that it is projective to claim that causal consequences of a physical theory relate to a putatively independent physical reality, as shown earlier in {21}.

However, when we choose to use an expression like ‘causal consequences’ we are reminded that we have left behind the primitive notion of causation and have replaced it with a purely *descriptive* but non-explanatory understanding. When a consistent pattern of regular functional dependencies has been observed, the expression of that pattern *can*, in principle, *only* be descriptive—if by ‘explanation’ we intend to mean, as we so often do, that we have found the genetic, productive, agency-based causes without which, in the counterfactual sense, their associated effects would not be.

To give another illustration: When it is said that the mass of a star “warps the space around it,” this must not be taken in the old putatively meaningful, conventional sense that the star somehow “produces” or “generates” a tighter metric of space. This point is important: It is not a matter of quibbling over the niceties of technical language, but rather a fundamental matter of the concepts that are presupposed, and of the understanding that we come to have on their basis. The amount of mass of a star *functionally defines* the degree of curvature of the space-time in its neighborhood. —There is no putatively meaningful “agency” involved, no “productive effort” that is being “exerted.” A star’s mass and its space-time curvature are *one and the same*: their *equation* expresses their functional interdependency.

While the primitive feeling of muscular push-and-pull effort associated by common sense with the succession of a cause and its effect may plausibly be considered a vestigial remnant of animism, it is significantly more than this: It is, as this chapter seeks in a variety of ways to emphasize, *metalogically projective*. In the above example, the very conditions of reference to a star’s gravity are, at present, established by the theory of general relativity; those conditions necessarily bring with them, as an unavoidable mathematical consequence of the current theoretical understanding of gravitation, an associated, functionally defined, curvature of space. Alternatively understood and expressed, gravity and metric of space (and of time) are, to repeat, functionally interdependent.

As we shall see later in {26} and {27}, both relativistic physics and quantum theory have largely been successful in relinquishing and then avoiding the traditional notion of causality. That they have accomplished this is an achievement that has clearly come about purely through the evolution of theoretical physics, not as a consequence of referential analysis as undertaken

here.

The result we have reached in connection with the concept of causality is an altogether different kind and conceptual level of result compared with the evolving concept of causality in physics, and the path by means of which we have reached this result has very evidently not in any way been based upon the method or the conclusions of physics. Although the de-projected concept of causality stands in substantial agreement with the current transformed and greatly diminished role of the concept of causality endorsed by many current physicists, the path in this study to that result has been solely through a series of conceptual analyses that show the conventional notion of causality to be metalogically self-undermining in a variety of ways. The resulting understanding of causality is independent of physics, as is its justification.

Hume dismissed the conventional notion of cause-and-effect because he claimed that the human mind cannot obtain an “impression” corresponding to the idea of “necessary connection,” and so the mind simply acquires a “habit” or “custom” due to the constant conjunction of putative “causes” and “effects” that have been experienced in the past. These are assertions for which strongly compelling proof has, ever since Hume’s time, been lacking—as witnessed by unrelenting philosophical controversies concerning conventional causality. The de-projected, revisionary concept of causality which we shall soon have reached is a result that cannot *not* be accepted without metalogical self-referential inconsistency.

### 23.4 The concept of causal network

We do not have a simple event *A* causally connected with a simple event *B*, but the whole background of the system in which the events occur is included in the concept [of causality], and is a vital part of it. If the system, including its past history, were different, the nature of the relation between *A* and *B* might change entirely. The causality concept is therefore a relative one, in that it involves the whole system in which the events take place.

– P. W. Bridgman (1961/1927, p. 83)

Earlier in this chapter mention was made of the dependence of the concept of cause-and-effect upon the recognition of *similarity* and upon the presupposed *selection* involved in identifying and discriminating a cause in relation to its effect. Presupposed by the identification of all cause-and-effect relations is the

need to be able to recognize that the selected cause and the selected effect bear certain similarity relations to one another sufficient to link the two in what we come to view as a relationship of functional dependency. An analysis of the preconditions that must be satisfied by such a recognition lies outside of the present study,<sup>276</sup> but even without such an analysis before us here, we are able to see that, in principle, the discrimination of cause and effect possesses its own Leibniz boundaries, which can often be continued and expanded to include an increasingly large scope of events that potentially are functionally related either to the identified cause, or to the effect, or to both.

The expression '*causal network*' will be convenient to refer to the fact that frequently our understanding of causal relations may be broadened to include other events, phenomena, processes, or other identifiable objects of reference that are judged to play causally contributory roles. Observations of individual cause-and-effect relations generally occur in a larger context, in terms of a background consisting not only of a record of past observations of such cause-and-effect relations, but of the more inclusive circumstances, states of affair, or conditions that formed the enviroing setting of the identified causes and effects. Within that context, an identified cause is recognized to be an indispensable contributing factor, but by itself it may not—potentially—comprise the totality of factors that may play a causal role in the cause's association with its identified effect.

The word 'potentially' is important. Insofar as a given cause-and-effect relation may be said to possess Leibniz boundaries, this implies an openness to additional observations which may identify, from among background events, phenomena, processes, or other identifiable objects of reference, those judged to have relations of functional dependency with respect to the cause and effect under consideration. This openness to additional, potentially relevant causal information may be taken in two senses. These parallel the two possible claims relating to the notion that the serial structure of temporal succession and of spatial order is "everywhere dense," discussed in {22.5}. As we saw there, it is metalogically self-undermining to claim that the continuability of a temporal series or the extendability of spatial observation relates to an "independently existing temporal or spatial continuum."

In much the same way here, the attentive reader will immediately see that it is projective to claim that such an implicit background of cause-and-effect relations possesses "ontological independence" with respect to reference frames in relation to which those relations are identifiable—in other words, that there purportedly exists a background of events autonomous of the

---

<sup>276</sup> For a detailed analysis, see the author's monograph, Bartlett (2015).

framework(s) in terms of which the observed causal relations are identifiable, out of which background “additional contributing causal factors” “in reality” play a role. We shall call these ‘*projections of hidden determinants*’. As we’ll see later, such projections frequently are at the basis of controversies in quantum theory; there, they appear in the form of what have come to be known as hidden variable theories.<sup>277</sup>

It is not difficult to avoid such projections. A non-projective openness to potentially relevant, additional causal information is clearly fundamental to scientific research. It is often possible to make additional observations which provide information relating to factors not previously recognized that play a role of functional dependency in identified cause-and-effect relations. We sometimes will wish to refer to potentially relevant causal factors that can be identified from the *implicit* background—to which, of course, reference must be possible—of observed cause-and-effect relations. On such occasions, we may associate the general notion of “causal networks” with an openness to the potential broadening of our understanding of causal relations.<sup>278</sup>

### 23.5 Determinism

All natural science is based upon the hypothesis of the complete causal connection of all events.

—Albert Einstein (2006/1910, p. 183)

Like its sister-topic causality, the subject of determinism has occupied philosophers for millennia. As we should expect, there are nearly as many notions of determinism as there have been philosophers. In the context of the present study, we have need only of a concisely stated definition of conventional, commonsense determinism.

---

<sup>277</sup> Such projections are related to another form of projection which I have called ‘projection of the implicit’ (Bartlett, 1974, 1975a), which is also associated, as noted in the previous chapter, with projections of temporal or spatial continuity.

<sup>278</sup> J. L. Mackie (1974, pp. 34-35) has used the expression ‘causal field’ to refer to the background causal context, claiming that this expression was introduced by John Anderson (1938). Anderson, however, did not employ this phrase in the 1938 paper cited by Mackie, while his use there of the term ‘field’ is vague, employed by him with a variety of meanings, and most often equated by him with ‘genus’.

In the preceding discussion of causality I have chosen to avoid the phrase ‘causal field’ in referring to the background causal context since using it in this way is not consistent with the general concept of field as employed in field theories of mathematics and physics. It is the latter concept that will later play an important part in this study.

The expression ‘causal networks’ was previously used, e.g., by Pearl (2009/2000), in its application to causal Bayesian networks.

The conventional notion of causality is an indispensable ingredient in the related commonsense notion of determinism which, like the notion of causality, involves a group of beliefs that are claimed to be meaningful and true. The conventional view of determinism consists of the following pair of beliefs, asserted together:

- ◆ All events are completely governed by previous causes, that is, every event is the result of antecedent causes which, if known, would predict that event with complete certainty.
- ◆ Every event is such that, had its cause or causes not occurred, the event would not have occurred.

The first belief was expressed by Einstein in the above quotation. He was thoughtfully cautious and referred to that belief as a hypothesis. In the commonsense notion of determinism, however, the assertion of the total causal control of present and future events by antecedent events is an outright, putatively meaningful assertion, not a mere hypothesis. There are two ways in which such a purportedly meaningful claim may be understood: One is to regard the claim as stating only an expectation, a hope, or a generalization based on past experience: the expectation that, as in past experience, any individual event could be predicted with certainty were its causal antecedents to become known. This is much like our everyday pattern of anticipation and behavior that follows suggestive rules which we follow in dealing, for example, with experienced three-dimensional objects (see {20.11}).

However, an alternative way in which the belief in universal determinism may be understood is as a putatively meaningful attempt to claim that there always exist causes that govern observed events, even when these causes are not accessible to us, and perhaps, in principle, cannot be. Here, we recognize, in slightly different clothing, a projection of framework-transcending causality, encountered earlier in this chapter; such a belief therefore undermines its own possible meaning.

The second commonsense belief associated with determinism—that every event is such that, had its cause or causes not occurred, the event would not have occurred—must similarly be rejected due to its projection of counterfactual causation.

The conclusion we therefore reach is that the conventional notion of determinism is metalogically self-undermining and hence without possible meaning. The exception to this result is when the conventional notion of determinism is solely an expression of the universal *hypothetical, speculative*

belief that we invest in the reliability of our *expectation* that events, as in our past experience, will be found to be associated with antecedent causes. Understood in this reduced sense, conventional determinism becomes a mere innocuous shadow, diminished to no more than the expression of a psychological predisposition to believe that what we have so far found to be the case will continue to be that way. Hume rightly called this a ‘custom’—that is, merely a ‘habit’.

### 23.6 Free will

Thoughts and intentions emerge from background causes of which we are unaware and over which we exert no conscious control.

– Sam Harris (2012, p. 5)

The everyday notion of free will embodies several of the projections we have already identified in connection with the commonsense notion of causality. We may take two approaches in analyzing the notion of free will: One, which is more direct and simplest, is, in the fashion of mathematics, to reduce it to a problem already solved. The other is to examine the notion of free will in some detail, and then find, unsurprisingly, that we reach the same conclusion.

The first approach is based on a result reached in {11.4}, that the negation of a projection remains a projection: that merely denying a projection is, like the projection itself, also metalogically self-undermining. The conventional notion of free will denies causal determinism insofar as human choices and volitions are concerned.<sup>279</sup> Those who deny causal determinism (that is, who assert its negation) as it applies to human choice and volitions will commonly claim, in disagreement with Harris in the introductory quotation, that human choices and volitions are *not* caused by “background causes of which we are unaware and over which we exert no conscious control.” —That is, they assert the negation of the determinist’s claim.

Such a putatively meaningful claim is, necessarily, devoid of meaning: We see immediately that the claim combines the *negation* of the *projection of*

---

<sup>279</sup> We take the occasion to note an early attempt to defend free will by means of self-referential argumentation by Boyle, Tollefsen, and Grisez (1976). The book’s use of pragmatical self-referential argumentation is unfortunately, and ironically, metalogically self-referentially inconsistent due to its reliance on the projective nature of the conventional notion of free will examined in the present chapter. This criticism of the argumentation by Boyle, Tollefsen, and Grisez was made in Bartlett (1979).

*genetic causation* (our choices are *not* causally produced by previous events); the *negation* of the *projection of counterfactual causation* (had events *not* been what they were, we would still be free to make our choices); the *negation* of the *projection of hidden determinants* (there were *no* antecedent causes that controlled our choices); and the *negation* of the *projection of framework-transcending causality* (there were *no* events of which we were unaware that caused our choices).

In short, encumbered as it is with all four projections of causality that we have identified, the everyday, purportedly meaningful, notion of free will is metalogically self-undermining and is therefore devoid of meaning.<sup>280</sup>

Belief in the freedom of the will comes close to being a human universal. It is the belief that human choices are generally free—unless, for example, an individual is said to be “under the (causal) influence” of drugs, intoxicants, mental disease, etc. The belief in free choice is unreflectively assumed to possess a meaning, and the belief is built on, layer upon layer, by the legal systems of many countries, and manifested in court proceedings, especially in the sentencing, punishment, and rehabilitation of those judged to have committed criminal offenses. If it can be shown that a murder was perpetrated by a defendant judged to have had his or her free will impaired at the time of the crime (e.g., thanks to a plea of insanity), he or she may escape execution, while the individual judged to have had free choice at the time of the crime may be escorted to the electric chair. Belief in free will pervades much of society; it is assumed unquestioning in the diagnosis and treatment of individuals by most clinical psychiatry and psychology; it is a basic credo of religions and a fundamental tenet of most systems of ethics and morality. The list could continue indefinitely. Free choice is nothing short of a widespread dogma and an ideology. Belief in it pervades the disciplines.

Where we do not find belief in free will, we often find a combined belief in free will under some conditions and determinism under others, and when neither of these beliefs is embraced, we most often find belief in the denial of free will, that is, determinism.<sup>281</sup> All of these alternatives are without possible

---

<sup>280</sup> *Mutatis mutandis*, the no-free-will claim made by Sam Harris and others is similarly, and necessarily, meaningless. One cannot meaningfully either affirm determinism *or* free will. Both are unavoidably projective.

<sup>281</sup> Interestingly, two surveys, one of evolutionary biologists and one of philosophers, led to the following results: Among the evolutionary biologists, 79% claimed to believe in free will, 14% denied it, and 7% left the question unanswered. (Graffin and Provine, 2007).

Among philosophers—certainly the main group of professionals who are alleged to have subjected the notion of free will to the greatest amount of sustained, careful, and critical analysis, and who might therefore be expected to have arrived at some unanimity—59.1% embraced “compatibilism” (an attempt to have it both ways—under some conditions, one is free; under



meaning because, as we've already seen, they entail projections: They undermine the very referential preconditions that would have to be met for them to possess a *possible* meaning.

Let us turn now to our second approach and consider the conventional notion of free will in greater detail. We recall that the commonsense notion of causality combines the following *beliefs*: A cause and its effect possess properties sufficiently in common to recognize that they are related; the cause usually occurs earlier in time than the effect; the cause is responsible for bringing about the effect; the occurrence of the effect is in some way dependent upon the cause; the relationship between cause and effect is one of necessity or indispensability in the counterfactual conditional sense that if the cause had not happened, the effect would not have occurred; the relation between cause and effect involves some kind of "tie," "agency," or the transmission of "productive power"; the cause and effect are regularly and uniformly associated with one another in that whenever the cause occurs the effect also occurs; and this pattern of uniformity and regularity is then relied upon in anticipating the likelihood of similar outcomes in the future. We are reminded that a number of these causal beliefs have been shown to be metalogically self-undermining.

When we consider the closely related commonsense notion of free will, the above beliefs continue to be relied upon and applied, but in negated form. Believers in free will wish to erect a partition to separate what they view as separate "worlds"—that of "external" causally connected events, in contradistinction to the "subjective, inner" world of choices and volitions. There are several links in a potential chain of events which believers in free will presume to be meaningful: Purportedly there are "outer" physical events that comprise the physical circumstances and background causes that potentially might influence an individual's "inner" choices and decisions. (I'll desist in using these scare quotes, which become wearisome but remain understood.) These choices and decisions are, however, believed to be independent of outer events and causes and to lead to what are called 'acts of will' or 'acts of volition' that are thought to bring about changes in that individual's inner thoughts or attitudes, and to control the person's outward behavior. Freedom of the will, in the commonsense view, consists in (a) denying the conventional notion of a causal relationship between outer causes and inner choices, thereby insulating and protecting the autonomy of inner decisions, combined with (b) asserting a conventional causal relationship between inner choices and subsequent acts of will or volition, as well as a second conventional

---

other conditions, not), 13.7% maintained the freedom of the will; 12.2% denied free will; with 14.9% in the ambiguous category of "other." (Bourget and Chalmers, 2013).

causal relationship between those acts of will or volition and subsequent behavior which those acts of will or volition bring about.

The first of these, (a), is, as we've already seen, the negation of an inherently projective notion of causality and is therefore devoid of sense. The second, (b), employs that very notion of causality, as it is believed to be translated to the inner subjective world of choice- and decision-making and their enactment in behavior; as we've shown, (b), too, is inherently projective. As long as the preconditions of reference cannot, in principle, be met, both the assertion and the denial of projective claims will necessarily be meaningless.

At the time of this writing it has become fashionable to believe that neurological research is relevant in answering the question whether the conventional notion of free will is true or false. From the standpoint of the present study, there are two responses to the widely shared high regard for the findings of brain science: One is to recognize that "the question whether the conventional notion of free will is true or false" is itself a meaningless question because it asks a question about a meaningless, projection-riddled notion, which, because it is meaningless, *cannot*, in principle, be either true or false.

A second response is to consider the abstract model presupposed by certain current neurological studies.<sup>282</sup> This is a model of brain activity that attempts to determine whether physiological or electrical brain activity does or does not temporally precede subjective inner choices and/or inner acts of will or volition. No matter what the results of such research show, they presume that the conventional notion of causality is meaningful: If antecedent physiological or electrical changes in the central nervous system allow the researcher to predict accurately what choice a test subject will make, the presumption is that this demonstrates that a conventional causal connection—as we have described this notion in some detail in this chapter—exists between those physiological or electrical changes and the individual's subsequent choice or act of volition. Such a presumption inevitably brings with it, as is now familiar to readers, a group of metalogically self-undermining claims. One can-

---

<sup>282</sup> Current neurological work in this area dates back at least to the recognition by German neuroscientists Hans Helmut Kornhuber and Lüder Deecke (1965) of the so-called human cortical "readiness potential" (*Bereitschaftspotential*) that is observable prior to a subject's initiation of a physical action. Years later, cognitive scientist Benjamin Libet and others were celebrated for their finding (Libet, et al., 1983; Libet, 1985) that this readiness potential is manifested about a half-second before an individual becomes aware of his or her intention to move.

Such empirical data are frequently believed to shed light on the free will/determinism issue, but, in principle, such data cannot accomplish this, since (as is shown in the text) the issue itself is metalogically self-undermining, destroying its own possible sense.

not—again, in principle—rationally expect neurological research to answer a self-undermining question.

### 23.7 Causality, determinism, and free will—in retrospect

In this chapter, we've examined various ways in which concepts and claims associated with the conventional notions of causality, determinism, and free will undermine their own preconditions of reference and, hence, of meaning. We've identified four varieties of causal projections: projections of counterfactual causation, projections of genetic causation (alternatively, projections of causal agency), projections of framework-transcending causality, and projections of hidden determinants. The dual purpose of the chapter has been to show the major ways in which the conventional notion of causality is meta-logically self-undermining, and then to describe how that putatively meaningful notion can be replaced with a revisionary, de-projected concept that is strongly compelling.

Our principal purpose as we progress in this critique of impure reason has been to distinguish the meaningful from the meaningless—specifically to identify and then to reject concepts and claims which transgress their meta-logical horizons, concepts and claims that attempt to extend beyond the boundaries which define the referentially forbidden. This task is critical and negative in the sense associated in earlier chapters {12, 13.5, 15.2} with the concept of a “negative science.”

The initial and most fundamental steps in developing a critique of impure reason are to recognize the unavoidable constraints upon possible reference, the boundaries of what is possibly meaningful, and then to identify and to reject concepts and claims which, while presupposing those constraints, undercut them and become meaningless.

If that were to be the end of our analysis, we would be left with a knowledge both of sets of strongly compelling injunctions and of a wide range of concepts and claims which violate them and which compel us, under pain of meta-logical self-referential inconsistency, to reject them. We would then be left without meaningful replacements for these rationally unacceptable concepts and claims.

Readers will recall from {15} that, to meet the need for such replacements, the task of the method of de-projection is completed through an attempt to provide a revisionary re-formulation of projective concepts and claims in a spirit of what I've called ‘respectful sympathy’: Its goal is to express in a manner that avoids projection what we sympathetically take, and at times must imagine, to be the pre-analytical intended “putative meaning” of

those concepts and claims. Such a re-formulation seeks to express the intended sense which, in reflective analysis, we imagine or suppose those theories, positions, and concepts *would need to have* in order for them to possess possible meaning. This final step of de-projection entails positive, constructive work that seeks to salvage what is capable of being salvaged from the results of the first step. As we've noted earlier, it is not always possible to accomplish this; sometimes nothing can be salvaged and a projective concept or claim must simply be rejected with no revisionary replacement.

Now that we have identified and rejected a group of metalogical projections involved in the conventional, commonsense notions of causality, determinism, and free will, what, if anything, may be salvaged?

With respect to causality, we've come to recognize that relations between pairs of events or between multiple sets of events may be judged to be causally related when the following conditions are met: Those events have come to be associated by virtue of certain properties they share which we select as relevant in their recognized causal pattern of association; these events have been observed regularly and uniformly to be correlated; the association between them can be expressed in terms of an invariant functional dependency; that functional dependency permits us reliably to predict one given the other; the functional dependency that we have formulated on the basis of the previous satisfied conditions is testable by means of falsifiable predictions; and, last, we leave an intentional blank space to be filled in by situationally determined requirements judged to constitute contributing factors.

For the reasons discussed earlier in this chapter, deliberately absent in this list of non-projective conditions is the commonly supposed condition of the temporal priority of causes in relation to their effects, a condition for which there is no compelling theoretical justification to include. As a result, from the standpoint of the metalogic of reference, causal relations are neutral with respect to temporal order.

When the above conditions together are met, we have grounds to claim that two events, or multiple sets of events, are causally related, leaving it open that future experience may show that such a judgment is mistaken or in need of revision. These conditions underscore the framework relativity of the de-projective concept of causality. They are conditions that, in principle as well as often in fact, can be satisfied. None of these conditions are undermined by an anthropologically and conceptually archaic, muscular push-and-pull, effortful notion of causality that insists upon the purported meaningfulness of genetic or agency-based causation, of counterfactual causation, of causation by hidden determinants, or of causation that is framework-transcending.

We've come to see that to understand the concept of causality in de-

projective terms is to remain entirely on a descriptive level. Cause-and-effect relations understood exclusively in the above de-projective terms provide us with an understanding of the functional dependencies and interdependencies of many kinds and classes of events, and frequently these functional relationships have important predictive value. Cause-and-effect relations understood de-projectively can be expected to meet our practical as well as theoretical needs since, understanding them as we have, they embody all that can be meaningfully salvaged. Where they fall short is in satisfying the hard-to-put-to-rest philosophical and psychological “compulsion” or “drive” discussed in {20.11}—the human impulse, desire, or urge to trespass beyond the boundaries of what is possibly meaningful.

With respect to the most commonly employed form of determinism, after we have applied the method of de-projection, little remains of what we may sympathetically imagine or assume to be its naive intended putative meaning. De-projected, determinism *can be* no more than the expectation, the hope, or a generalization based on past experience—the expectation or hope that, as in the past, it may be possible to predict an arbitrarily chosen event with certainty *were* its causal antecedents to become known. Determinism, so understood, is nothing more than a speculative hypothesis, one which attempts to be so all-inclusive that it can only remain hypothetical. When belief in determinism attempts to push beyond these boundaries it undermines itself and becomes meaningless.

Finally, with respect to the commonsense notion of free will, once divested of its projective elements, freedom of the will, as in the case of determinism, must be relinquished as yet another futile attempt to transgress beyond the boundaries of the possibly meaningful. Perhaps the proponent of free choice successfully will be able to cast doubt upon possible causal relations (here, I of course mean causal relations as de-projectively understood) between, for example, certain identifiable brain signals and temporally subsequent conscious choices or volitions to move a part of one’s body—even though such causal relations may be expressed by brain scientists in the form of well-confirmed functional dependencies. But even if successful in casting such doubt, the upholder of free will still faces an implacable set of metalogically unavoidable facts—that it is projective to claim any of the following: to assert that our choices are *not* causally produced by antecedent events; or to claim that had events *not* been what they were, we would still be free to make our choices; or to claim that *no* previous events controlled the choices we have made; or to claim that there were *no* events of which we were unconscious that caused our choices. Once we have recognized the projective nature of these claims and realize the in-principle meaninglessness of the projective

beliefs tied to them that so many people continue earnestly and sometimes fervently to endorse, the conceptually primitive notion of free will should be allowed to die, receive a long-overdue burial, and then be consigned to oblivion.

## Projections of the Self and of Solipsism

Among the perennial problems of philosophy are those of the self and of solipsism. The first problem has traditionally led, directly or indirectly, to the conflict between realism and idealism, while idealism in its extreme has often taken the form of radical subjectivism. In the limit, idealist-subjectivism, in turn, can take the form of solipsism, which has challenged on an epistemologically fundamental level the existence of an independent external world and the existence of other minds. The concepts of the self and the philosophical standpoint of solipsism have been at the basis of some of the most highly debated questions for which philosophers through the ages have sought to provide answers. The fact that none of the answers so far proposed has compelled the unified rational assent of philosophers as a group attests, one may plausibly believe, either to the intractable degree of the conceptual difficulty of the problems themselves, or perhaps even to a willful blindness endogenous to the philosophical profession (discussed in {2}). But there is still another and more persuasive alternative, as this chapter seeks instead to show. This last alternative claims that the perpetual controversies over the conflicting answers that have been proposed by generations of philosophers to the problems of the self and of solipsism are evidence of underlying conceptually confusions which, once disentangled, bring about the complete dissolution of these problems—in a manner that cannot *not* be accepted without metalogical self-referential inconsistency. Once we have reached this result, we shall be in a position to develop further the concept introduced earlier in {5} of non-relational, agentless reference.

### 24.1 The self

As in the preceding chapter concerned with the notions of causality, determinism, and free will, it will continue to be my purpose here to advance the objectives of a “negative science” by identifying a group of ways in which the notion of the self, as commonly used by philosophers and non-philosophers alike, is metalogically self-undermining. Having done this, I turn attention to a de-projective revisionary re-formulation which attempts to salvage from the

traditional notion of the self that which, in principle, is meaningful.

The traditional, commonsense notion of the self may be broken down into a set of claims or beliefs that have served—sometimes singly, in individual variations, and sometimes in combinations—as the main basis for philosophical disputation about the self. Those claims include the following:

- (1) Thinking entails a thinker—that is, there are always two components of thought: a subject (the self) and objects of which the subject is aware.
- (2) The self is an existing entity.
- (3) The self is the center of experience of the world.
- (4) The self is a bearer or owner of its states.
- (5) The self possesses or is characterized by faculties which it exercises in a wide variety of ways.
- (6) The self is an agent, the cause of thinking—that is, the processes of consciousness result from the activity of the self.
- (7) Every experience is had by a self, by an at least implicit “I”; alternatively, consciousness by the self is a universal characteristic of experience.
- (8) The sum total of the self’s experience comprises what metaphorically may be called ‘a container’ or a ‘phenomenal field’—that is, the self’s consciousness holds or encloses all that it experiences.
- (9) The self serves as a limit which the individual cannot ever go beyond or get away from: this is the position of solipsism.

We shall examine each of these claims in turn.

#### **24.2 Thinking entails a thinker—that is, there are always two components of thought: a subject (the self) and objects of which the subject is aware**

“Something is thought, therefore there is something that thinks”: this is what Descartes’ argument amounts to. [T]hat there must be something “that thinks” when we think, is merely a formulation of a grammatical custom which sets an agent to every action. In short, a metaphysico-logical postu-



late is already put forward here—and it is not merely *an ascertainment of fact...*

– Friedrich Nietzsche (1909/1901, §484)

These impressionistic and aphoristic pronouncements by Nietzsche, potentially insightful but not epistemologically compelling, have been asserted and sometimes defended by a number of philosophers, going back more than a century to von Schubert-Soldern (1884, pp. 65ff), Mach (1886, pp. 19ff), Driesch (1923/1912, p. 19), Husserl (1913, pp. 65, 160, and *passim*), Aster (1913, p. 33), Schlick (1925/1918, pp. 147f), Natorp (1923/1910, pp. 41ff), Russell (1921, p. 18), Carnap (1967/1928, pp. 105ff), and others since then.

Common sense, however, has become unreflectively habituated to the post-Cartesian split between mind and matter, between subject and object, and has become so much accustomed to their separation that to question it seems not only unnatural but contrary to what has come to feel intuitively self-evident. The subject-predicate structure of English and of similar languages contributes, as Nietzsche commented, to the promotion of the same unquestioned bias. As we shall see in subsequent sections, this bias permeates and is presumed by most of the other claims distinguished above that pertain to the notion of the self.

Among branches of philosophy, the branch most directly relevant to claims relating to the self has been phenomenology due to its descriptive and analytical focus upon the structure of experience, where presumably one might most reasonably expect to find evidence of the self. Early in the development of phenomenology, Franz Brentano (1924-25/1874) developed the notion of intentionality, the view that consciousness inevitably involves relatedness to objects. A great deal of effort has been expended by phenomenologists in efforts to shed light on and to critique the central notion of intentionality, which many believe describes much of the experiential basis for the subject-object relation and for the notion of the self and its alleged role in conscious life.

Our primary interest in this work is, however, not phenomenology, but rather a study of the preconditions of reference. In this chapter, it will be our objective to identify those preconditions of reference that would need to be satisfied—if possible—in order for any of the nine claims relating to the notion of the self to meet our standard of meaningfulness.

Consider the first claim—that thinking entails a thinker, that is, that there are always two components of thought, a subject (the self) and objects of which the subject is aware. We need to ask, What conditions would, of neces-

sity, need to be satisfied in order for this claim to be capable of referring as it putatively suggests? Let us suppose that a given reference frame permits reference to temporally successive experiences  $e_1$ ,  $e_2$ ,  $e_3$ , etc. Let us further suppose that each temporally successive experience retains a record of the previous experience(s), so that  $e_2$  includes a memory, or other record, of  $e_1$ ;  $e_3$  includes a memory of  $e_2$  and  $e_1$ ; etc. We further assume that it is possible for certain experiences—those of a special kind which we shall call ‘reflective’—to involve, in some of various specific forms which here do not concern us, awareness of “a subject” or “a self,” which we’ll designate by ‘ $s$ ’. In these terms, consider the following succession of experiences:  $e_1$ ,  $e_2(e_1s)$ ,  $e_3(e_2e_1s)$ , etc. Here ‘ $e_2(e_1s)$ ’ refers to the experience temporally subsequent to  $e_1$  which is such that  $e_2$  retains a record of  $e_1$  conjoined with the association of  $e_1$  with an experience, in some particular form, of “a self”  $s$ . The temporally subsequent  $e_3(e_2e_1s)$  is the experience that makes it possible to refer to the experienced fact that  $s$  was experienced *in retrospect* as having been involved in the initial experience in this nested series. (Often such a claim asserts that “ $s$  was implicit in the initial experience.”<sup>283</sup>)

This example has been constructed to represent the following situation: The experience  $e_1$ , as it was initially experienced, does not explicitly involve  $s$ , but  $s$  was noticed only by means of a later experience  $e_2$  when, in retrospect, the association of  $s$  with  $e_1$  was experienced. In other words,  $e_1$  comprises a “bare,” “subjectless,” or “subject-neutral” experience; it is an experience in which there is no “consciousness of self.” Such experiences are quite common and have, as we shall see later, been recognized explicitly by many philosophers.

### 24.3 Temporal preconditions of reflective reference: An aside

The preceding example can be instructive in several ways: First, as a semi-formal representation of common, routine experience, it helps to direct our attention to the temporally successive and retentive capacity of much familiar experience. Second, it supplies a model-in-miniature of what has come to be called ‘reflective awareness’: We recognize that it is by virtue of the retentive capacity of experience that we are, in principle, able to take note of experienced changes. The change which the reflective awareness that is enabled by  $e_3(e_2e_1s)$  is a function of a reference frame’s ability to compare and contrast the earlier experiences  $e_1$  and  $e_2(e_1s)$ , and by so doing, to recognize that, in retrospect,  $e_1$  is recalled as having involved the experience of  $s$ . From the

---

<sup>283</sup> For more a more detailed analysis of the concept of the implicit, see Bartlett (1974, 1975a).

standpoint of this study's focus upon preconditions of reference, "reflective thought," "reflective ability," and the concept of conceptual "reflection" as it can be considered generally, all of these, we should now recognize, require, as conditions of their possibility, a reference frame's capacity to retain temporally successive states, as illustrated in the foregoing abridged example. Without the capacity to retain temporally antecedent states, reflection becomes impossible.

Third, the example can be instructive by making clear that reflective experience, understood as we have, has unavoidable, necessary metalogical horizons which, if transgressed, lead to self-undermining projection. We immediately see this in the present case once we recognize that, in principle, there is no possible way in which the subjectless experience  $e_1$  can be claimed *to have involved* awareness of  $s$  prior to and independently of temporally later experiences  $e_2(e_1s)$  and  $e_3(e_2e_1s)$ . Expressed differently, to claim that every experience entails a subject who experiences, that every experience involves a self which has that experience, is to make an in-principle meaningless claim, one that is metalogically self-undermining. Such a claim would require, in order to be possibly meaningful, that it should be possible, in this example, to experience that which, in principle, cannot be experienced. The very preconditions of reference that must be satisfied in order to refer to a subject or self that is associated with any given experience cannot, in principle, provide a basis for the claim that an experience  $s$  was involved under conditions that exclude reference to  $s$ . To make such a claim is to assert a metalogically self-undermining projection.

#### **24.4 Projections of reflection and projections that thinking entails a thinker**

We have just encountered two different additional varieties of projection which we shall from now on call '*projections of reflection*' and '*projections that thinking entails a thinker*'. We recognize projections of reflection when we realize that the results of reflective analysis cannot meaningfully be applied retroactively to pre-reflective experience. We recognize projections that thinking entails a thinker when we realize, not only that not all experience involves awareness of a subject, but that it is devoid of meaning to claim that all experience is experience had by a subject. As we continue to consider the other naive claims relating to the self listed toward the beginning of this chapter, the importance and consequences of these projections will become progressively clear.

### 24.5 The self is an existing entity

*Subject I* this is the term we apply to our belief in an *entity* underlying all the different moments of the most intense sensations of reality; we regard this belief as the effect of a cause....

– Friedrich Nietzsche (1909/1901, §485)

Reference to “*an entity*” most commonly means that some “*thing*” is claimed to be an identifiable and re-identifiable object. In the case of physical objects, as we’ve previously seen, their Leibniz boundaries express our expectations that additional information about them is likely to be forthcoming as we experience them as a function of other perspectives. We experience this kind of expectation that further information will potentially come to our attention when many varieties of objects of reference are in view, not only those that are physical. Whether the objects in question are objects of fiction, or abstract objects, or physical things, to regard them as “entities” is most commonly to claim that, *relative to information that is presumed could be provided by alternative perspectives*, we anticipate that we potentially may obtain more than the given data with which we are presented at a particular time ({21}). It is important that we bear clearly in mind that the preceding claim relates to our *expectation* that supplementary information may potentially become available, and nothing more than this.

When the entities under consideration are physical objects, we saw that to assert their independence while withholding the basis for possible reference to them is to transgress a metalogical horizon, that is, to engage in projections of the external world ({21.1}).

We further saw that sensory information relating to other people is also experienced in perspectival form, information that is usually given to us in a temporally and spatially discontinuous manner, but sufficient for us to recognize uniformities and regularities fundamental to the capacity to identify and re-identify individual persons and to distinguish among different persons. As with physical objects, however, we also saw that when a purportedly meaningful additional step is made, one that seeks to transgress beyond the horizon of the individual’s possible reference frame, there result projections of other minds and/or projections of other minds as things-in-themselves ({21.3}).

In the context of our present referential analysis of the notion of the self as an existing entity, the same observations and conclusions apply: To the extent that it is possible to refer to information upon which one’s personal identity is based—information which, again, is given in a fragmentary

perspectival way and is experientially discontinuous in time— to that extent the individual is capable of recognizing his or her own identity and is able to re-identify at subsequent times that essentially the same set of data judged to be salient is in view. One's identity as such a "self" is based on many such kinds of data, including sets of memories and sensations, physical appearance, abilities, dispositions, etc. We shall equate an individual's identity with the class of all such data.<sup>284</sup> No matter how long and how detailed such a listing or accumulation of individualizing identifiable information becomes, and no matter how it may be anticipated that this listing potentially may become ever longer, it has been a compelling temptation on the part of Cartesian common sense and on the part of many philosophers to wish to go beyond such a framework-relative inventory, and to claim of "the self," as in the cases of the "external world" and of "other minds," that "more than this exists."

When one falls victim to referentially delusional thinking of this kind, one becomes a casualty of what I shall call '*projections of the self as an existing entity*'. We find such projections widespread in psychological theorizing that wishes to postulate the existence—"above," "behind," or "underlying" experience—of entities such as the "ego" or "superego," or when philosophers similarly assert the existence of a "transcendental ego" or a "transcendental unity of apperception," or their cousins. Characterized as they have been in a wide variety of ways by many different authors, such purportedly meaningful notions that a self, ego, superego, transcendental ego, etc., exist as framework-autonomous entities violate the referential constraints upon possible reference and result in claims devoid of possible meaning.

#### 24.6 The self is the center of experience of the world

It is of the essence of knowledge that it is in the first person....  
Actually given experience is given in the first person; and reality as it is known in any case of actual knowledge can be nothing, finally, but a first-person construction from data given in the first person.

– C. I. Lewis (1934, pp. 127-128)

[I]f one is reconciled to the inevitability of describing the world from himself as center, a unity is thereby automatically

---

<sup>284</sup> From the very different standpoint of his method of "logical construction of the world," Carnap gave a similar definition: "The *self* is the class of autopsychological states" (Carnap, 1967/1928, p. 241).

restored to the world, the unity conferred by the necessity of seeing everything from a single origin. This is not the illusory unity which we formerly thought it had, but is the only unity we can use, the only unity we need, and the only unity possible in the light of the way things are.

– P. W. Bridgman (1959, p. 248)

In earlier chapters we have seen how the identifiability of objects of no matter what kind is framework-relative. We have also noted that often such objects are given in perspectival form. The “self,” as this concept is gradually gaining content in our analysis, is no different: Like any object of reference, reference to an individual’s identity is framework-relative, while the class of information in terms of which that identity comes to be recognized is, as we’ve noted, also given in perspectival form.

The above quotations make the claim that the individual subject, in experiencing the world and attaining knowledge of it, does so, in Lewis’s terms, “in the first person,” thereby establishing, as Bridgman expressed this, a “center,” “a single origin,” from the standpoint of which the world comes to be understood. These claims are of a very different sort than the framework-relative results re-stated in the preceding paragraph. The claims made by Lewis and Bridgman make the identity of the individual a reference frame in and of itself. Not only is this to claim something very different about the self than the way we have come to understand it, it is to make an inherently projective claim that can, in principle, have no meaning. Certainly it is true that individual observers have identities: We can refer to distinguishable classes of memories and sensations, physical characteristics, abilities, dispositions, etc., which enable us to identify individual persons, but it is quite another matter to claim that any one of these classes, which are themselves objects to which reference is possible, comprises a reference frame from the standpoint of which such a class can be regarded as “the center” or “a single origin” of experience and of knowledge.

The difference I wish to underscore is, to engage in a bit of metaphor, analogous to the shift from the Ptolemaic understanding of the physical universe to the Einsteinian: Where Ptolemy judged the Earth to be the center of the universe, we might, in a parallel but poetic translation, think of Einstein’s approach as leading to the recognition that the notions of “center” and of “the universe” are themselves ideas whose meanings are consequences of, and wholly a function of, the adoption of a methodologically explicit framework-relative understanding.

To state matters more carefully: When one says, as did Lewis, that “experience is given in the first person,” one can only mean that experience *includes* information that serves to identify an individual: This information is a “content” of experience, but it is not, and cannot meaningfully be said to be, as we shall see in a few pages, a “container” of that experience. When one claims, as Bridgman did, that it is inevitable for the world to be described with the observer’s “self” as center, one can only mean that an observer’s reference frame includes information about him or her (whether this is a matter of “inevitability” we shall consider shortly).

Reference to an individual’s identity is assured when it is possible to refer to a particular class of memories and sensations, physical characteristics, abilities, dispositions, etc., which provide the basis for recognizing that individual’s identity. Reference to such a class, however, requires an appropriate frame of reference: That class is not the frame, but is itself identifiable only as a function of an appropriate framework of reference. When an attempt is made purportedly to “endow” a group of data with a “stand-alone” existence as we see in Lewis and Bridgman, a projective delusion is promoted whose referential preconditions, in principle, cannot be satisfied.

The putatively meaningful notion that everything experienced and known by an observer forms a perspective “brought about” by his or her (or its) identity is, as we shall see, a confused combination of several additional projections. The variety we have been discussing in this section we shall call ‘*projections of the self as center of experience*’.

#### 24.7 The self is a bearer or owner of its states

The commonsense notion of the self often involves the claim that the self is the “bearer of its states”—in other words, that the self, in some allegedly meaningful fashion, “supports its consciousness” in the capacity of the original meaning of the word ‘bearer’, that is, as “one who carries or sustains,” or in the words of the *King James Bible* (2 Chron. ii. 18), individual selves serve as “the bearers of burdens.” This easily pictured characterization of the “bearer” as one who *carries*, who is the *porter* of physical or emotional *loads*, has been applied by common sense to the identity of individual persons, and then the “bearer” has been reified: The commonsense self has come to be accepted as a real, existing *carrier* of the contents of awareness.

Very much in parallel with this naive use of picturesque-figurative language, the self has been portrayed as “the holder,” “the owner,” or “the possessor” of its states. Philosophers, such as P. F. Strawson (1959, pp. 88ff and *passim*), have expended thought and time in analyzing such seemingly inno-

cent, purportedly meaningful, but, from the standpoint of the present study, fundamentally incoherent modes of thought and expression.

It is, of course, one thing to dismiss as naive and unwarranted the notions of the self as “bearer of its states” or of the self “as owner” of its experiences, but it is quite another to show that the referential preconditions that would need to be met in order for such notions to have possible meaning cannot, in principle, be satisfied. In the preceding sections of this chapter we have already done this: We have recognized four projections of the self thus far: projections of reflection, projections that thinking entails a thinker, projections of the self as an existing entity, and projections of the self as center of experience. In reviewing the definitions of these projections, the reader will immediately see that the putatively meaningful notions of the self as the bearer or the owner of its states incorporate all four of these metalogically self-undermining claims; we shall call this combined projection, ‘*projections of the self as bearer or owner of its states*’. The notions of the self as the bearer or owner of its states are not only misplaced and confused, but, more importantly, they undermine their own possible meaning and are to be rejected from rational consideration.

#### **24.8 The self possesses or is characterized by faculties which it exercises in a wide variety of ways**

Here again, but in a different manifestation, we encounter archaic attempts to characterize personal identity in reifying terms; we can easily see how these undermine their own possibility of meaning:

The commonsense notion of “faculty” is a remnant of the “faculty psychology” that was in vogue during the latter part of the 19th century.<sup>285</sup> A “faculty” was considered to be the source of a variety of “powers of the mind” such as memory, the will, reason, etc. Psychologists who reified these abilities, aptitudes, or functions claimed that their causal sources were corresponding existing mental “faculties” which, when actively exercised, bring about the expression of those abilities, aptitudes, or functions. In other words, for example, the exercise of the faculty of reason was thought to cause mental activity of a certain kind, and to produce “acts of reason.”

Referential preconditions that would have to be met by claims that a self exists that possesses such “faculties” which it “exercises” in a variety of “acts,” cannot, in principle, be met: Although we can refer identifyingly, for example, to memories, and, in so doing, presuppose an appropriate reference

---

<sup>285</sup> Cf., e.g., Wundt (1890), James (1890), Flügel (1951/1933).



frame in terms of which such reference is possible, the purportedly meaningful postulation of a causal source of those memories—an alleged “faculty of memory” claimed to exist beyond that reference frame—attempts to transgress beyond the metalogical horizon of such a reference frame, exceeding its possible referential capacity and hence its possible meaning.

The conventional and commonsense notion that the self is the causal basis of “acts” and “activity” of various sorts, to which we now turn, is intimately linked to the foregoing projections which we shall call ‘*projections of mental faculties*’.

#### **24.9 The self is an agent, the cause of thinking—that is, the processes of consciousness result from the activity of the self**

The “mind,” *something that thinks*, at times, even, “the mind absolute and pure”—this concept is a ... result of false introspection, which believes in “thinking”: in the first place an act is imagined here which does not really occur at all, i.e. “thinking”; and, *secondly*, a subject-substratum is imagined in which every process of this thinking has its origin, and nothing else—that is to say, *both the action and the agent are fanciful....*

Owing to the phenomenon “thought,” the ego is taken for granted; but up to the present everybody believed, like the people, that there was something unconditionally certain in the notion “I think,” and that by analogy with our understanding of all other causal reactions this “I” was the given *cause* of the thinking....

– Nietzsche (1909, §§477, 483, original italics)

The occurrence of the content of a thought constitutes the occurrence of the thought. Empirically, I cannot discover anything corresponding to the supposed act, and theoretically I cannot see that it is indispensable.... [T]he act in thinking is not empirically discoverable, or logically deducible from what we can observe.

– Bertrand Russell (1921, pp. 16, 18)

Philosophers have searched for the “acts” “caused by the self,” but the search, at least for many philosophers, has been futile; they have been unable to find

evidence to substantiate the notion that “activities” are brought about by “an underlying self.” Carnap therefore observed: “Philosophical introspection [*Selbstbesinnung*] has led philosophers of various persuasions to the same result, namely, that the original processes of consciousness must not be envisaged as the activities of an acting subject, the ‘self’” (Carnap, 1967/ 1926, p. 261).

Here, once again, it is important to recognize the significant difference between, on the one hand, blunt *assertions* that the “activities of an acting subject” are neither empirically discoverable nor logically deducible from what is observable, and, on the other hand, a *proof* that the very notion of such “activities” is metalogically self-undermining. The former outright and unproved claims leave open the potential that they are entirely misplaced or are fundamentally subject to revision, but the latter closes the door to the possibility that such “activities” can have possible meaning.

The results we have reached in previous chapters are cumulative: A number of these results apply to the putatively meaningful notion that “the self is an agent, the *cause* of thinking.” We’ve identified three varieties of projection that relate to the conventional notion of causality, all of which apply to the notion before us: (*a*) projections of counterfactual causation, (*b*) of genetic causation (alternatively called ‘projections of causal agency’), and (*c*) projections of hidden determinants. In the present chapter, we’ve so far identified an additional six projections that relate to the notion of the self: (*d*) projections of reflection, (*e*) projections that thinking entails a thinker, (*f*) projections of the self as an existing entity, (*g*) projections of the self as center of experience, (*h*) projections of the self as bearer or owner of its states, and (*i*) projections of mental faculties. As we summarize and review these nine results, it is not hard to see how, together, they render the notion of “the self as causal agent” rationally unacceptable and fit to be tossed on the scrapheap of objectionable conceptual debris inherited from the past.

The conventional and commonsensical notion that there are activities of the self which causally result in the processes and manifestations of consciousness characteristically involves a group of claims purported to be meaningful: that conscious states entail the existence of a self, a subject, which is conscious of them (*e, f*); that “behind the scenes” of conscious awareness there exists a causally productive, genetic self (*b, c, d*); and that this self is the center of experience, the bearer of states of awareness, and the source of “mental faculties” (*g, h, i*). —These, combined, I suggest, make up the typical, conventional, commonsensical notion of the “activity” that is believed to be involved in conscious (and, some think, also in “unconscious”) life.

If the foregoing analysis is accepted—which on pain of metalogically self-undermining inconsistency it must be—the combined force of the nine projections (*a – i*) strongly compels us rationally to *reject* the notion of “acts that are involved in thinking and in other mental processes” possesses *possible* meaning. We shall call this combined projection by the short name, the ‘*projection of agency*’.

**24.10 Every experience is had by a self, by an at least implicit “I”—that is, consciousness by the self is a universal characteristic of experience**

As I have elsewhere,<sup>286</sup> I shall call the above claim ‘*the spectator theory*’: It is the view that awareness by an “I” or “self” is a universal characteristic of experience.

A first observation concerning the spectator theory is that it must be formulated from the standpoint of reflective thought: To claim (*a*) that *implicit* in experience is a self, an “I,” is a claim that can only be made from the standpoint of a nested series of temporally retentive states as described earlier in this chapter, while to make the universal generalization (*b*) that consciousness by the self characterizes all that is experienced, is a second reflective claim. The first of these claims conventionally intends to assert the autonomous existence of an “I” that putatively “stands apart” from experience in a relationship of observer to the observed. The second claim also typically asserts this, or it may simply be an expression of a habitual expectation formed from an experienced regularity of the presence in past experience of members of the class of memories and sensations, physical characteristics, abilities, dispositions, etc., which provide the basis for recognizing the identity of an individual.

In claim (*a*) we immediately recognize the occurrence of both a projection of reflection and a projection of the self as a existing entity, and this is also the case in the claim typically made by (*b*). Claim (*a*) seeks to apply the results of reflective thought retroactively to pre-reflective experience, and intends to assert the existence of an “I” or “self” as a framework-autonomous entity, existing “above and beyond” what is experienced. As we have seen, these are both horizon-transgressing attempts, which undermine their own possible sense. Such attempts are clearly believed to be meaningful by those who engage in them, victimized as they are by their own propensities to violate the metalogical horizons of their presupposed reference frames.

---

<sup>286</sup> Bartlett (1970, §§2.6, 2.7).

We shall call projections of this variety ‘*projections of spectator consciousness*’.

**24.11 The sum total of the self’s experience comprises what metaphorically may be called a ‘container’ or a ‘phenomenal field’—that is, the self’s consciousness holds or encloses all that it experiences**

We shall call this view ‘*the myth of consciousness as a container*’:<sup>287</sup> It is the notion that all experiences that include awareness of the self—as we have come to understand this concept in terms of members of an individual’s identity-defining class of memories and sensations, physical characteristics, abilities, dispositions, etc.—make up a “container” that “holds” all of that which an individual experiences. Let us call this identity-defining class of experiences ‘*M*’. The myth of consciousness as a container may then be expressed in the form “all experience (of the individual identified by *M*) is contained in *M*.”

This purportedly meaningful notion that such self-associated consciousness forms a “container” or a “receptacle” is common and widespread, and is encouraged by such often-used phrases as “it was part of *my* experience,” “*my* experience contains all that I am and all that I’ve come to know,” etc. This view has come to play the role of a favored myth both in philosophy and in common parlance and thought. It has also, when expressed in its most extreme form, led to the philosophical position of solipsism, which we shall examine in the next section. Let us take a moment to ask what the referential preconditions must be in order for the notion of a “container of consciousness” possibly to make sense.

We notice, first of all, that what I’ve called a ‘container of consciousness’ is believed, by those who believe in the myth, to “hold” or to “enclose” all that is experienced. This is to take a sub-set of experiences and claim that all an individual experiences is contained within that subset. This is to give priority by means of the notion of “containment” to self-associated experience, and to diminish or disregard what I shall call ‘self-neutral’ experience, which is discussed later in this chapter.

When we ask what referential preconditions must be satisfied in order to make this claim of containment, we should need to be able to refer to a relation of containment obtaining between all experiences (of the individual identified by *M*) and *M*. This, the individual is incapable, in principle, of doing.

---

<sup>287</sup> Also discussed in Bartlett (1970, pp. 216, 265ff).

Possible reference by the individual identified by *M* to members of *M* requires a frame of reference supplied by the experience of the individual identified by *M*. From the standpoint of that reference frame, reference cannot, in principle, obtain to a relation of containment between all his or her experiences and members of *M*, which are themselves experiences of the self defined by *M*. We shall call projections of this kind '*projections of consciousness as a container*'.

But such projections do not exhaust the confusions at the basis of the myth of consciousness as a container. To see them in greater relief, let us consider the philosophical position of solipsism.

#### **24.12 The self serves as a limit which the individual cannot ever go beyond or get away from: The position of solipsism**

[O]ne cannot study the idealistic philosophers long without coming in touch with solipsism. Idealism is always trembling on the brink of it. The philosophers shy away from it like startled fawns, but their arguments continue to lead them back to it and so far as I can judge they escape it only because they will not pursue them to the end.... It has a completeness and an elegance that make it infinitely attractive. Since I cannot suppose that everyone ... will know all about the various philosophical systems, the instructed reader will perhaps forgive me if I state briefly what solipsism is. The solipsist believes only in himself and his experience. He creates the world as a theatre of his activity, and the world he creates consists of himself and his thoughts and feelings; and beyond that nothing has being. Everything knowable, every fact of experience, is an idea in his mind, and without his mind does not exist. There is no possibility and no necessity for him to postulate anything outside himself.... It is a perfect theory; it has but one defect; it is unbelievable.

– W. Somerset Maugham (1950, pp. 648-649)

This paragraph by Maugham tells us, in a nutshell, what solipsism was understood to mean to a writer of fiction who devoted serious study to philosophy. Maugham concluded that solipsism is a "perfect theory," but one that is unbelievable—except perhaps by a solipsist. Certainly it is true, as Maugham

noted, that philosophers tend to shy away from solipsism; I think it is safe to say that many philosophers believe the challenge it poses is inordinately hard to counter, even by kicking a stone as Dr. Johnson did.

Solipsism takes belief in the alleged limitative boundaries of a postulated self to its extreme. The position may be defined as the repudiation by an individual that “there are objects outside of that individual’s experience.” Understood in these terms, it can be considered a variation of the myth of consciousness as a container, which in the preceding section we’ve already recognized to be metalogically self-undermining. The denial by solipsism that “there are objects outside of the individual’s experience” is also projective since such a denial asserts the negation of the projective claim of realism (see {21.5}).

It is rare for a philosopher seriously to promote solipsism. Philosopher-physicist P. W. Bridgman was, in a sense, one of these. We recall that earlier in this chapter I quoted his suggestion that one should be “reconciled to the inevitability of describing the world from himself as center.” Already this is a step toward solipsism. I would like to give Bridgman the opportunity here to supply us with a more complete explanation, since solipsism is so rarely advocated with anything approaching sincerity:

In the last analysis science is only my private science, art is my private art, religion my private religion, etc. The fact that in deciding what shall be my private science I find it profitable to consider only those aspects of my direct experience in which my fellow beings act in a particular way cannot obscure the essential fact that it is mine and naught else. “Public Science” is a particular kind of the science of private individuals...

This position, which I suppose is the solipsist position, is often felt to be absurd and contrary to common sense. How, it is asked, can there be agreement as to experience unless there are external things which both you and I perceive? Part of the hostility to the solipsist position is, I think, merely due to confusion of thinking, and there is a strong element of the pseudo-problem mixed up here. If I say that an external thing is merely a part of my direct experience to which I find that you react in certain ways, what more is there to be said, or indeed what other operational meaning can be attached to the concept of an external thing? It seems to me that as I have stated it, the solipsist position, if indeed this be the solipsist

position, is a simple statement of what direct observation gives me, and we have got to adjust our thinking so that it will not seem repugnant. (Bridgman, 1936, pp. 13, 15)

As we reflect on Bridgman's point of view we take note of several claims he has woven together: There is the claim that the experience upon which the scientist bases his findings is equated with a range of experience that Bridgman identifies as experience which he judges to be that of a subject who "has" that experience; here, Bridgman once again, as quoted in {24.6}, suggests the notion that the scientist's "self" is the "center" from which he describes the world. We saw that this notion itself involves the metalogically self-undermining projection of the self as center of experience. We could, on these grounds alone, dismiss Bridgman's sympathetic acceptance of solipsism. And yet, as we shall see a little later, there may be something salvageable and important in what Bridgman sought to communicate.

For now, however, we've already quickly identified three projections involved in the conventionally understood position of solipsism: the projective negation of the claim of realism (equivalent to the projection of idealism), the projection of the myth of consciousness as a container, and the projection of the self as center of experience. (If we wished for a more thorough inventory, other projections, too, are involved; they include: the projection of reflection, the projection that thinking entails a thinker, the projection that the self is an existing entity, and plausibly others as well.) That multiple projections are involved informs us that solipsism is, in its traditional meaning, rationally unacceptable and to be rejected.

There is, nevertheless, another aspect of this analysis that requires our attention.

### **24.13 The neutral character of pre-reflective experience**

By this section heading I mean that much pre-reflective experience is neutral with respect to its association with members of an individual's identity-defining class of memories and sensations, physical characteristics, abilities, dispositions, etc. That is, experience is often given without any reference to what we have come to understand as a self. This is an empirical fact about experience, familiar to many who are untutored in philosophy and attested to by many of the tutored. Moritz Schlick stated this observation in terms that are clear and direct:

The strongest emphasis should be laid on the fact that primitive experience is absolutely neutral or, as Wittgenstein has

occasionally put it, that immediate data “have no owner”.... To see that primitive experience is *not* first-person experience seems to me to be one of the most important steps which philosophy must take towards the clarification of its deepest problems. (Schlick, 1936, p. 359)

Long before Schlick’s time, German physicist, philosopher, and satirist Georg Christoph Lichtenberg (1742–1799) made much the same observation:<sup>288</sup>

We know only the existence of our sensations, representations, and thoughts. *It thinks*, we should say, just as one says, *it lightnings*. [*Es denkt, sollte man sagen, so wie man sagt: es blitzt.*] To say *cogito* is already too much if we translate it as *I think*. (Lichtenberg, 2012, *Notebook K*, ¶76, p. 152)

Schlick was strongly influenced by Lichtenberg’s few suggestive fragments relating to the neutral character of much experience. Schlick went on to say:

The words ‘I’ and ‘my’, if we use them according to the solipsist’s prescription, are absolutely empty, mere adornments of speech. There would be no difference of meaning between the three expressions, ‘I feel my pain’; ‘I feel pain’; and ‘there is pain’. Lichtenberg, the wonderful eighteenth-century physicist and philosopher, declared that Descartes had no right to start his philosophy with the proposition ‘I think’, instead of saying ‘it thinks’. (Schlick, 1936, p. 365; see also Schlick, 1925/1918, pp. 147f).

Carnap made the same observation, referring to both Schlick and Lichtenberg; Carnap accordingly introduced the appropriate phrase ‘the subjectless given’ (Carnap, 1967/1928, p. 106). Bridgman similarly claimed: “It is almost the

---

<sup>288</sup> Early thinkers who reached the same conclusion about experience include those referenced in {24.4}; other philosophers include Avenarius (1907-1908/1888) and Volkelt (1918, pp. 59ff). There is of course an ancient and much longer tradition of Eastern practice that has led to the same result; to quote but one sample passage that makes this evident:

[I]n one technique of jana yoga, the practice which employs a refinement of discriminating intellect, one may minutely observe one’s experience, noting that no component of experience contains a “self.” This is the path of “*neti, neti*” (I am not this, not this), in which one progressively detaches the self-sense from the elements of experience with which it has formerly identified. (Walsh and Shapiro, 1983, p. 426)



same thing for me to say 'There is awareness' as to say 'I am conscious' ” (Bridgman, 1959, p. 224).

This “self-neutral” character of experience, although it plays a frequent part of everyday experience, is often overlooked or neglected in philosophical discussions of the self, and especially this is so in connection with the position of solipsism.

Given that solipsism, as traditionally understood, is untenable because it falls victim to multiple projections, we may wish to ask how solipsism may be understood in a different light when one takes into account the self-neutral character of much experience.

The point at which we have so far arrived in this chapter rejects the conventional commonsense notions that “thinking entails a thinker”; that “the self is an existing entity”; that it is “the center of experience, the bearer, or owner of its states”; that it is “the source of mental faculties”; that “the self acts or has a role in experience as an agent”; that “the self possesses a framework-independent existence as an “‘I’ ‘above and beyond’ what is experienced”; and that “there is a self that functions as the ‘container’ of experience.” In keeping with the goals of negative science, these results all point to ways in which the conventional, commonsense notion of the self violates constraints upon possible reference and goes astray by making claims devoid of possible sense.

Once these traditional claims are swept aside, we are left with a de-projected residuum that may be expressed as follows:

There are many varieties of objects of reference for which correspondingly appropriate frames of reference assure the possibility of reference. Among these there are certain kinds of objects of reference that pertain to what we have come to understand as the self: They consist, as we have seen, of sets of memories and sensations, abilities, dispositions, etc., which together comprise members of an individual’s identity-defining class of characteristics. A class of such characteristics comprises objects of reference which are familiar to us, but it is, in principle, devoid of meaning to claim of such objects of reference, as is so often done, that they point to or form the basis of or reveal in any sense “a self that stands apart from and beyond what is experienced.”

We are, to be sure, unaccustomed to thinking of experience in these “minimalist” terms, unencumbered by the delusional sleights-of-hand which our habitual but metalogically self-undermining concepts seduce us into accepting at their putatively meaningful face value.

Earlier in this work, a good deal of attention was given to the concepts of framework relativity and of self-enclosure. We have recognized a series of related claims, which include the following:

- philosophical systems often are “closed upon themselves” much like ideological systems of belief, as, in the extreme, can also be the systems of belief of delusional psychiatric patients;
- from a phenomenological standpoint, experience forms a self-enclosed “field,” beyond which it is not possible to go;
- the reflexive, vertical, non-ordinal frame of reference established by the metalogic of reference forms a self-enclosed system;
- the mathematical and the quantum-theoretical concepts of field may in their respective ways be understood to exemplify self-enclosure;
- the “referential field” established by a given frame of reference may also be characterized as self-enclosed, and in a particular manner that has both significant ontological consequences and the consequences that come from understanding a referential field’s metalogical horizon, beyond which reference and meaning are impossible; and
- the method of de-projection itself has a self-enclosed structure that has “no outside.”

These observations relating to framework relativity and self-enclosure yield the following result when applied to the conventional, commonsense notion of the self which this chapter has been considering in some detail. The de-projected salvageable understanding that we have come to may be expressed in the following fashion:

Frames of reference have certain metalogical horizons, a recognition of whose constraints leads, as we’ve seen, to an acknowledgment of the projections of the external world and of other minds, and to an acceptance that both realism and idealism are both metalogically self-undermining. Furthermore, a framework of reference cannot be “assigned to a subject”—as if a subject who is identifiable by means of an appropriate set of memories and sensations, abilities, dispositions, etc., somehow, in some purportedly meaningful way, “contains” or is “the center of experience” of that framework. Again, such an identity-characterizing set contains objects of reference, but such a set of

objects of reference cannot of itself, on pain of projection, be thought of as a “container” of experience, as an independently existing “spectator,” or causal source of psychic “activity.” A frame of reference is not equivalent to or contained by a self; the latter is only identifiable by means of a suitable frame of reference. We have come to recognize these things, but they bear repeating given the stubborn tendency of old ways of thinking to hold us in their grip.

The metalogical horizon of a reference frame can inform us of the parameters of constraint upon what is relative to that frame while exhibiting the property of self-enclosure; in this specialized sense, we reach what might be called a ‘*solipsism without a self*’—a phrase that is no more than an intentionally odd, ostensibly paradoxical, “back door” way of acknowledging both framework relativity and the self-enclosure of the metalogic of reference.

To bring this admittedly elusive idea into better definition, I’m reminded of one of Wittgenstein’s many artfully shrouded mystifying pronouncements:

The subject does not belong to the world; rather it is the limit of the world.

Where *in* the world is a metaphysical subject to be found?

You will say that this is exactly like the case of the eye and the visual field. But really you do *not* see the eye.

And nothing *in the visual field* allows you to infer that it is seen by an eye....

Here it can be seen that solipsism, when its implications are followed out strictly, coincides with pure realism. The self of solipsism shrinks to a point without extension, and there remains the reality co-ordinated with it. (Wittgenstein, 1961/1921, §§5.632–5.64)

These fragmented paragraphs, typical of Wittgenstein’s chosen enigmatic and aphoristic style of writing, are often quoted, but they have not, to my knowledge, been shown to mean clearly and unambiguously any one definitive thing. Perhaps we can associate a definitive meaning with these bits and pieces, or at least a few of them, a meaning that may or may not have occurred to Wittgenstein. —Certainly, a great advantage of writing obscurely and in epigrams is that interpretation is open to all, no one can be proved wrong, and a greater number of unemployed philosophers can find jobs.

Readers who have followed the steps taken in this study will not, on the one hand, be able to accept the meaningfulness of much that Wittgenstein claimed in the above quotation (e.g., that “the subject ... is the limit of the world,” or that “solipsism strictly carried out coincides with pure realism”),

but, on the other hand, perceptive readers may perhaps appreciate the added meaning we may attach to Wittgenstein's last sentence, "The self of solipsism shrinks to a point without extension, and there remains the reality co-ordinated with it." The meaning we attribute is this:

In our analysis, the notion of the "self" has indeed "shrunk away"—not to a "point without extension," but to only one, among many, of the varied kinds of objects of reference to which reference is possible. Having undergone this "shrinkage," experience has become divested of one of its main delusionary distortions. We are left not with "reality" as it has traditionally been construed, but with what I have earlier called 'referential fields', which we shall explore in greater detail in the next chapter. As we shall find later on, referential fields provide a basis for objectivity—but an understanding of objectivity freed, however, from the projective notions of realism, idealism, other minds, the self, and other extraneous notions from which the concept of objectivity has had a difficult time extricating itself.

...

It admittedly can be a laborious task, demanding patience and fortitude, to rid a wide range of concepts—as well as our habitual processes of thought that employ them—of self-undermining phantasms left to us from a more primitive, naive, and gullible past, a past inhabited by a people who easily invested their credulity in unobservable spirits and specters which have frequently, much like invasive species, found their ways into the structure and content of ordinary language. Both common sense and much philosophy have unfortunately fallen heirs, uncritically, to these conceptually delusional attempts to stake out claims devoid of meaning beyond the horizons of their own presupposed frames of reference.

## Non-relational, Agentless Reference and Referential Fields

Much of the focus of this book has been the nature of framework relativity and the consequences that follow from its recognition. Early chapters have made the claim that the concept of reference has inappropriately and mistakenly been burdened by extraneous delusion-inducing notions that are inherently projective. In {5.4}, I asked for the reader's interim cooperation in deliberately setting to one side his or her habitual association of reference with its associated conventional activity-based conception. We recall that the latter is the commonsense notion that reference is something which a mind in its thinking, or which a person in his or her attempts to communicate, *does*. What he or she *does* is allegedly to *direct*, in a causally productive way, the agent's own attention, as well as the attention of a hearer if one is involved, to that to which reference is, in this allegedly active process, *made*. However, {5} may, even to the assiduous and thorough reader, seem like long ago, so a few words will be useful as a reminder of a promise made early in this long study:

Many pages back I made the claim that by divesting the concept of reference from its subject-object, activity-based conception we would come to find ourselves in an advantageous position to understand the parameters of constraint which this study of the metalogic of horizons of meaning and possibility seeks to describe and apply. It has become increasingly clear in the course of these pages that many of our epistemologically centered concepts, as well as our thinking—which often employs those very concepts as our preferred conceptual vocabulary—have become thoroughly permeated by metalogically self-undermining beliefs and claims that express those beliefs.

Up to this point, we have developed the ideas of this study using a provisional—some might call it a ‘temporarily emasculated’—notion of reference that has asked, as we've noted, for the reader's partnership, putting to one side, at least for a time, the habitual, naive, traditional, conventional, commonsensical referrer-referent, activity-based conception of reference. It should now be clear to the reader precisely why that request was made: The preceding two chapters, dealing with projections of causality and those of the

self, have shown that conventional causal claims and claims that involve a putative self in the role of agent cannot, in principle, be meaningful. It can therefore make no sense to incorporate such causal and agent-based claims in the concept of reference. The request made of the reader to place the causal and agent-based notions of reference to one side has had this justification.

The “minimalist” concept of reference that has therefore been employed throughout this book may, for some comparatively rare readers, *suffice*—that is, the minimal concept may be adequate for their purposes, for their conceptual and practical needs. However, given the widespread compelling human propensity that we’ve now encountered numerous times, to “wish for more” and to “go beyond limitations,” for the majority of readers it is likely that a broader, explicitly integrated account of reference is likely to be desirable, but one which self-consciously conforms to the referential constraints that come with an acceptance of framework relativity and an awareness of the inevitable limitations of metalogical horizons. This chapter is intended to meet those readers’ needs.

### 25.1 Non-relational, agentless reference

As a theory of reference formulated on the level of maximum theoretical generality whose prime concern is with the preconditions of reference, the metalogic of reference has involved a number of specialized, central, and theoretically abstract concepts.<sup>289</sup> As described in depth in previous chapters, the function of these fundamental concepts is to clarify and to provide the reflective means to analyze the parameters of constraint that underlie the possibility of reference. It will be useful at this stage to take note of certain essential interrelations among three of these central concepts that have had major roles in this work.

The concept of *frame of reference* is of course one of these. We recall that a frame of reference provides a basis for the identifiability and re-identifiability of one or more sets of objects, and in these terms, and on a conceptually basic level, the meaning of any concept, claim, or question relies upon, or more accurately is a function of, an appropriate associated frame of reference. In earlier chapters, we have discussed in some detail how, in providing a basis for the identifiability of a range of objects, frames of reference function in a coordinative capacity that is at the very basis of possible knowledge and the communicability of that knowledge.

---

<sup>289</sup> I intentionally refrain from calling these ‘constructs’, a term which itself suggests an activity-based genesis.

In this context, it is important to underscore the realization we have come to that frames of reference cannot in any sense be “reified”: Emphasis has been placed on the relational dynamic that intimately ties possible objects of reference to the coordinative frameworks in terms of which they are identifiable. We have seen how any object of reference has ingredient in it, as an integral constituent of its identity, the constitutive structure of the frame of reference in terms of which it is identifiable.

A second central concept, that of *embedment*, was therefore introduced to describe how, “embedded” in the identity of an object of reference is the interdependency relation between the object and one or more appropriate reference frames. Far from “reifying” frames of reference, we have come to understand them as dynamic systems, an interdependent function of which the sets of objects to which they may refer are relative. This understanding is at the core of our concept of framework relativity. It is especially important to stress this dynamic interdependency, given the human tendency to think of a “framework of reference” as—in some purportedly meaningful way—separable, detachable, or autonomous with respect to that framework’s domain of possible objects of reference. “Built into” the identity of an object of reference *is* its very framework relativity.

The concept of frame of reference and that of embedment are at the basis of the *non-relational approach to reference* to which the metalogic of reference leads. We recall ({5.6}) that by a “non-relational” conception of reference is meant an understanding that reference is fundamentally identity-specification or identity-recognition in the sense that an object and its identity are not disjunct, but essentially fused: *To be is to be an object of reference relative to an appropriate framework*. The object, its very identity, is indissolubly linked to the reference frame(s) that provide the basis for its identifiability. The nature of that link is informative: It tells us in what sense the object *is*. To be an object *is* to have a specifiable identity ({18}).

The conventional subject-object, referrer-referent, activity-based notions of reference as a “voluntary,” “choice-driven,” “causal process” that is “initiated by a self,” “the referring agent,” all, as we have seen, lead to metalogically self-undermining claims which, in principle, *can have* no meaning. Many philosophers and non-philosophers alike have developed a deeply rooted habituation in accepting and asserting these and similar claims. Such habits have clearly been buttressed by the subject-object, referrer-referent, activity-based grammatical structure of English and similar languages. For readers committed to the rational rules of admissibility described in {17}, the naive notion of reference strongly compels its rejection and replacement by its de-projective revision.

The naive notions of reference and of the “activity of referring” that is conventionally believed to be associated with reference, presume, as we recall from {5}, the purported meaningfulness of a cluster of projective notions: These notions relate to the “mental states of the agent” and to “the agent’s volitional acts and purposive activity,” which, in turn, are believed to be the “causal expressions of the agent’s mental abilities or faculties,” and whose exercise in “acts of referring” succeed or fail in “directing the attention of other minds” to the “agent’s intended object(s) of reference.” Having shown that each of these expressions, which have here been placed in scare quotes, is metalogically self-undermining, it is clear that, if we are to develop a philosophically neutral, maximally general understanding of reference, we are forced to abandon the incoherence of the naive notion of reference, and to replace it with a conceptually sound understanding.

Our gradually refined understanding of the concept of reference requires revision that is clearly radical in nature—‘radical’ both in the sense of ‘extreme’ and in the sense of ‘basic’. The detected presence of projections in the naive view of reference leaves us with a de-projected concept of reference that is both non-relational and agentless. Once that which is metalogically self-undermining is “stripped away,” what can be salvaged is an understanding of reference that is not only non-relational and agentless, but can most clearly be understood in terms of a third central concept introduced in earlier chapters; we called that concept ‘*referential fields*’.

## 25.2 Referential fields

The concept of referential fields provides a theoretically integrated understanding of the intertwined pair of concepts we’ve just summarized, that of frameworks of reference and of embedment. Previous pages have emphasized that possession of an identity and the possibility of reference are “intimately linked” in the relational dynamic which itself “intimately ties” possible objects of reference to the frameworks in terms of which they can be identified. Suggestive phrases like those just placed in quotes can now be explained in a more informative way. The concept of referential field provides a clearer and more explicit explanatory understanding of the special variety of relational dynamics with which we are concerned.

In earlier chapters, the general concept of field has been described, particularly as it has come to be employed in mathematics and in physics ({8.3.2}). But it is specifically in theoretical physics that this concept has been developed in ways that help to shed light on the concept of referential



fields, which are at once both significantly different from the fields studied in physics, and yet which share certain similar fundamental properties.

<b>THE CONCEPT OF FIELD IN PHYSICS</b>	<b>THE CONCEPT OF REFERENTIAL FIELD</b>
1. A field defines a relationship of functional interdependency between the states of elements within the field and the positions which those elements may have within the field	1'. A field defines a relationship of functional interdependency between the identities of a set of possible objects of reference within the field and an associated reference frame
2. The properties of such elements determine the degree to which they are affected by the field	2'. The embedded constitutive structure of objects of reference in the above set determines their in-principle identifiability as a function of the field
3. Changes in the state of such elements are understood in terms of their interaction with the field in which they are present	3'. Specification of an object of reference involves a dynamic inter-relationship between the field and the object, i.e., possible properties of an object of reference are functionally defined by its referential field
4. An "empty" field, one without elements present, has only the status of characterizing <i>potential</i> interactions with elements, were they to be introduced	4'. A referential field considered without explicit reference to objects in that field is only a <i>potential</i> basis for a set of in-principle identifiable objects
5. A field is such that with each position in the field can be associated a vector of potential force: a force, that is, which has magnitude and direction	5'. A referential field satisfying conditions 1'.-4.' is such that coordinate positions in the field can be associated with relationship to a potential observer, discussed later in this chapter.

**Table 25.1 Parallels and differences: The field concept in physics and the concept of referential fields**

Since the time of Newton in the 17th century physicists have sought to understand the three principal forms of action: action-by-contact (also called ‘action-by-impact’), action-at-a-distance, and action in a continuous medium. Especially in connection with the study of magnetic force and the effects of gravity, during the 20th century the concept of field has proved to be especially fruitful. Action in the continuous media of fluids had already come to be well-understood: The flow of fluid and the properties of elastic media provided physicists with parallels that seemed to fit the continuous nature of magnetic force and of gravitational effects, which had not been thought to exemplify either action-by-contact or action-at-a-distance. The gradually emerging understanding of action in a continuous medium made it possible to characterize variations of magnetic force or of gravitational conditions at different points in space without the assumption of remote causes. The force experienced by an object placed in a magnetic field or the conditions experienced by an object in a gravitational field could then be functionally understood in terms of the interdependency of the field, the properties of the object, and its location in the field.

This form of understanding in terms of *functional interdependency* has become very familiar to us in this study—not primarily in connection with the phenomena studied by physics, but in connection with the preconditions of reference. In physics, the study of action in a continuous medium has evolved to a point, as noted in {10.2}, that subatomic particles may be understood as excitations of specific fields, while matter may now perhaps be more adequately investigated in terms of the properties of the Higgs field. At the time of this writing, the major shift from a corpuscular to a continuous field model characterizes a very substantial portion of ongoing research in theoretical physics, both on the quantum and on the cosmological level.

It is a shift which, in a parallel manner, we have made in recognizing that *embedded* or *ingredient* in the very identity of objects of reference are the coordinating principles which render their identification possible. When the identity of an object has been referred to as an “instantiation” or “realization” of the referential field of its presupposed reference frame, the concept of referential field serves, in purely conceptual terms, to shift what we recognize as *meaningful* from the object—which conventionally is often considered *per se*, as though it could be identified in some framework-independent manner—to a metalogical level of understanding on which we recognize the functional interdependency of objects and their presupposed reference frames. The concept of referential fields makes this interdependency explicit. Let us see how this is the case.

The field concept in physics is typically thought to express the conceptu-

ally fundamental claim that there is a functional interdependency between the value of a field at any point and the properties or states of an object present in that field at that point. This interdependency is often, misleadingly, called the ‘field effect’, for one of the main objectives of field theory has been to avoid the introduction of traditional mechanical cause-and-effect notions. In addition to the unfortunate choice of terminology, the field effect is often not stated precisely: It is not that a field expresses a *de facto* functional interdependency; it is rather that a *description* or *formulation* of a given field informs us of the functional interdependency between the field at any point and the *potential* properties or states of an object that *may* be present at that point. — To give an example: With respect to a gravitational field, the function  $V$  is called the gravitational ‘potential function’ not because it is intended to describe a specific property of the field independently of any masses present in the field, but rather because it literally describes the *potential* gravitational force that *would be* experienced by any mass—if one were present—in the field at that point. The force involved is, however, called ‘gravitational potential energy’, a phrase that employs the word ‘potential’ in a distinctly different second meaning, i.e., in its physical, empirically quantifiable *non-modal* sense. The field concept, nonetheless, is essentially a *modal* concept, a fact often lost sight of when physicists are engaged in calculating the conceptually concrete gravitational potential energy of objects.

The italicized terms ‘description’ (or ‘formulation’), ‘potential’, and ‘may’ in the preceding paragraph are of paramount significance. The “field effect” is fundamentally a *description* of a formally envisaged range of *possibilities* under a specified set of constraints. *This is the core meaning of the field effect in physics; it is also the central meaning of the concept of referential fields.*

We recall that a frame of reference may be described in terms of an associated range of *possible* objects of reference; the theoretically general concept of possibility reached in {7} understands possibility as a function of a framework’s parameters of constraint. The concept of referential field has played a central, at times implicit, role in the dynamic concept of system-based identity and identification developed in this work. In previous chapters,<sup>290</sup> we have come to understand a referential field as a highly general structural concept relating to ordered wholes, that is, relating to events or ensembles of objects that are recognized to form totalities, and whose constituent objects of reference must—on pain of metalogical self-referential inconsistency—be understood in their relationships of functional interdependency with respect to the frames of reference that make their identifiability possible.

---

<sup>290</sup> {10, 14, 15, 20, 21, and 24}.

In this sense, the concept of referential fields provides an understanding of the functionally interdependent nature of the organization, the structural order, of integrated totalities that form dynamic systems of reference. In recognizing the field-based nature of systems of reference, we have come—repeatedly, in example after example in this third section of our study—to take note of the metalogically self-undermining nature of attempts to endow a wide variety of objects with horizon-transgressing autonomy. In rejecting such attempts, we’ve come to recognize the strongly compelling rational necessity to comply with the interactional principles of organization that constitute the parameters of constraint upon which organized systems of reference and meaning depend for their very possibility.

### 25.3 Referential fields and the place of the observer

The role of the observer in the metalogic of reference may, for many readers, appear by this point to be severely compromised or even eliminated: Our rejection in the previous chapter of nine projections relating to the self would, on the surface, appear to remove “*the observer*”—to whom the commonsense notion of the self is routinely applied—from any meaningful place in our study. But to draw this conclusion would be hasty. It is important to recall that in developing the concept of frame of reference as understood in this study, a place was included in systems of referential coordination that may link together, for example, spatial locations with objects, times with events, objects and times with observers, etc. ({10.1}). Following this, the reader will remember that in the provisional and abstractly simplified formalized description of referential consistency as a criterion of meaning ({11.4}), identifying reference was expressed in the familiar conventional form of a ternary relation among a person, an object, and a set of space-time coordinates; such a person may of course serve the function of an observer. The role of the observer was also noted in describing the Leibniz boundaries of a referential field, which potentially can be extended indefinitely to accommodate, for example, changes in the observer’s location or perspective ({14.2}). The potential involvement of the observer was also recognized in our discussion of the deprojective analysis of theories as systems of coordination that interrelate the following: possible objects of reference, a formal or other system that provides a contextual background, and perhaps including reference to a time, spatial position, and/or relationship to one or more observers, or recording instruments of detection, measurement, calculation, etc. ({15.1}). Later, our analysis of “external perceptual space” was understood as the imagined or conceptualized comprehensive reference frame in terms of which the per-

spective spaces of individual observers are considered to form parts ({21–22}) and in terms of which “the same object” may be recognized by multiple observers ({22.1}). We have also recognized that separate observers do, of course, have identities which involve distinguishable classes of memories and sensations, physical characteristics, abilities, dispositions, etc., all of which, sometimes separately, often in combination, enable us to identify individual persons ({24.5}). —In short, the potential role of the observer or observers has been affirmed at various stages of this study. And yet, how, in light of the various self-undermining projections of the self, are we to understand the concept of the observer and the observer’s place in a metalogical account of reference?

We have reached a de-projective level of understanding in terms of which individual persons—which in the present chapter we shall include in the class observers—are identifiable, as just noted, by virtue of their distinguishable memories and sensations, physical characteristics, dispositions, etc. That we identify individuals in such ways as these is commonplace, although, as we saw in the last chapter, the commonplace way of conceptualizing this fact typically involves the cluster of projective claims about the self which we’ve been forced to reject.

Both the concept of the observer and the identifiability of individual observers can only be understood relative to frameworks of reference in terms of which individual observers *may themselves be objects of reference*. The concept of the observer in our study has the meaning of any member of a certain special class of objects to which reference is possible: We’ll call that class the ‘class of observers’. Observers, as we shall understand them, fulfill a number of important and distinguishable purposes: First and foremost—in terms of the referential functions they serve—they are capable of retaining a record of past events; in this capacity, they have the ability to retain information in a temporally serial manner. Fulfilling this function is not necessarily to be limited to the human capabilities of memory; as we’ve noted, that function may be supplied by non-human recording devices of detection, measurement, calculation, etc. Observers, then, need not be human. Whether human or non-human, a minimum condition that observers must satisfy is the ability to retain information concerning past states. Satisfaction of this condition, as we’ve seen in previous chapters ({10, 12, 22}), is required in order that a framework be capable of re-identifying objects of reference, in order for reflection to be possible, and in order, as well, to make detection of error possible.

Observers also function as identifying markers; they tell us the source of observed information, and often also when and where that information was

obtained, and under what conditions and circumstances. Such “markers” should be understood in terms of the concepts introduced earlier in {5.7} of coordinates and of the coordinative function of frames of reference.

Specific and multiple varieties of frames of reference may be associated with individual observers. For example, a given observer may possess any combination of such well-known abilities as visual, auditory, tactile, etc.; measurement, calculation, or probability-estimating abilities; pattern recognition, problem-solving, understanding, or judgment abilities, etc. An observer may be especially well-equipped to detect subatomic or astronomical phenomena, to make predictions of certain kinds, etc. All of these abilities are associated with corresponding frames of reference.

Unfortunately, the everyday notion of the observer expresses the naive and projective beliefs that with each observer is associated his or her own individual “frame of reference,” and frequently, moreover, that this “frame of reference” is to be equated with “the individual’s consciousness or experience.” In commonsense belief, as we’ve seen, the observer’s “frame of reference” is then thought to serve as a “container” of “his or her awareness and experience,” a notion we’ve come to reject as a self-undermining projection of the container of consciousness, while, also according to common sense, the observer’s “self” is thought to comprise the “center” of that experience, an instance of the projection of the self as center of experience. From the standpoint of the metalogic of reference, frames of reference are not, in any possibly meaningful sense, equivalent to or “contained” by selves; the latter are only identifiable relative to and in terms of such reference frames.

In short, the concept of observer here retains much of its ordinary meaning, but it does not, however, suggest that the frame(s) of reference associated with an observer is/are “equivalent to,” or are “contained in,” the observer; the observer is only identifiable as a complex object of reference from the standpoint of a reference frame that associates the observer’s identifying properties with information provided by the observer. The observer, in other words, may have a place *in terms of a referential field*, but in no possibly meaningful sense can it be said that such a referential field is “attributable to,” “created by,” or “originates in” the observer.<sup>291</sup>

The above-mentioned association of the observer’s identifying properties and of information the observer provides is important: As we’ve stressed, observers are themselves identifiable objects. The referential field in terms of which an observer’s observations are formulated and understood provides the basis for possible reference not only to the observer but to information reported by observer. Stripped of the conceptually delusional notion of other

---

<sup>291</sup> See the related discussion in {19} of projections of invention.

minds, observers are simply particular kinds of objects of reference among others.

In the preceding table of parallels and differences relating to the field concept in physics and the concept of referential fields, the fifth condition, 5', stated that coordinate positions in a referential field can be associated with their potential relation to an observer. In commonsense terms, we often situate information in terms, for example, of time, place, and person; here, from the standpoint of the metalogic of reference, much the same claim is made, but on a modal level: A referential field is itself, as we've seen, an essentially *modal* concept, defining a relationship of functional interdependency between the identities of *possible* objects of reference within the field and an associated reference frame, in a manner such that *possible* properties of these objects are functionally defined by their referential field. We've also stressed that an observer is only identifiable as an object of reference from the standpoint of a reference frame that associates the observer's identifying properties with information obtained from the observer. In other words, it is the referential field—in terms of which observer and information associated with the observer are given—that makes it possible, under suitable circumstances, to associate coordinate positions in that field with a relation to an observer.

Information may be linked to individual observers and their reported observations in much the same way that information is linked to specifiable times and spatial locations. In such cases, coordinates for these times and locations can, in principle, be provided; among such coordinates can then be included specification of any relevant observer(s). However, as emphasized in this and the preceding chapter, reference, as we've come to understand it in conceptually neutral terms, is both non-relational and agentless; it follows that *not all referential fields include an association with an observer*. In other words, a given referential field is not necessarily associated with a set of an observer's individualizing properties, which in the human case often include physical characteristics of a particular body, set of dispositions, etc.

Referential fields are modal fields in terms of which the independency relations among objects and their frameworks of reference can be recognized. The observer may have a place in such fields—and, as we shall see in the next chapter, the observer may in certain specialized fields have a *necessary* place—but the observer's "standing" or "status" in such fields is that of an object of reference—not as a "center of experience" or "activity-based source of consciousness."

### **25.4 Referential fields and metalogical horizons: A brief summation**

The concept of referential fields has been introduced in order to make explicit in this study of the preconditions of reference the functional interdependency of objects and the frameworks of reference they presuppose. We've seen that the referential field associated with any given reference frame describes, literally, a *field of possibilities*, the scope of that reference frame's possible objects of reference. The purpose served by such a description is to make explicit that such a field is potentially extendable and therefore possesses no delimited boundaries, and is in that sense unbounded, and yet to show that it exhibits what I've called 'reactive boundaries' that become evident when attempts are made to trespass beyond a referential field's metalogical horizon. We've seen that the concept of referential fields is a modal concept whose purpose is to make explicit the range of referential possibility under the set of constraints determined by any given frame of reference, and by doing this, to enable us to recognize the boundaries of reference and of meaning that inevitably accompany our presupposed frames of reference.

In the earlier discussion of the metalogic of reference in relation to transcendental argumentation ( $\{9\}$ ), the distinction was made between the "positive" constraints that function as prescriptive rules that must be followed relative to a given frame of reference, and the "negative" constraints that function as a reference frame's boundaries of admissible reference. The concept of self-enclosure was developed as it relates both to the self-enclosure of the reflexive, maximally general theory of the metalogic of reference, and to the self-enclosure of specific frameworks of reference ( $\{4, 10\}$ ). We then described the ontological implications of the self-enclosure of referential fields, appropriately termed 'relativistically self-enclosed' ( $\{10, 18, 26\}$ ). We next reached an epistemologically neutral, de-projected understanding of reference as both non-relational and agentless, and recognized the place of the observer in referential fields (this chapter and the preceding).

From the standpoint of the resulting metalogical field-based analysis of reference, we have repeatedly found in connection with one major philosophical problem after another that putatively meaningful but necessarily delusional attempts are made to disconnect a wide variety of objects of reference from their essential framework relativity. We have found that the dynamic, modal concept of referential fields is both a necessary consequence of, and itself involves a recognition that entails an understanding of, what meaningfully may be said to be.



## Relativity Physics and Quantum Theory: Preamble

In relativity theory, the emphasis on the dependence of all phenomena on the reference frame opened quite new ways of tracing general physical laws of unparalleled scope. In quantum theory, ... the logical comprehension of hitherto unsuspected fundamental regularities governing atomic phenomena has demanded the recognition that no sharp separation can be made between an independent behaviour of the objects and their interaction with the measuring instruments which define the reference frame.

– Niels Bohr (1959/1949, p. 224)

Of the multitude of theories that have been proposed across the entire range of disciplines, from the exact sciences to the humanities, there are none that exemplify the degree of deliberate, self-conscious, critical analysis concerning framework relativity as do Einstein's two theories of relativity and quantum theory. In fundamentally different ways, as we shall see, relativity theory<sup>292</sup> and quantum theory exemplify rigorous, disciplined adherence to theoretically based constraints of framework relativity. Relativity theory and quantum theory provide instructive examples of ways in which certain of the fundamental principles that define the heuristic of the metalogic of reference appear almost as though they have unselfconsciously, implicitly, and intuitively guided the pioneering physicists in both areas of study. And, as we shall see, in different ways, relativity physics and quantum theory bear out certain of the principal results to which the metalogic of reference leads.

Both relativity theory and quantum theory comprise specialized theoretical systems of identification. They are, in other words, abstract frameworks that establish groups of rules in terms of which certain sets of physical objects can be identified. The sets of objects and the referential preconditions that are presupposed differ among the special theory of relativity, the general theory,

---

<sup>292</sup> In the subsequent chapters relating to physics, when I use this expression, both the special and general theories are included.

and quantum theory, as we shall discuss in the following pages.

Both relativity physics and quantum theory are unusual among theories developed in no matter what discipline because both, although in very different ways, go to the very heart of framework-relative analysis as it applies to the phenomena that make up their objects of study. Both conduct the reflective analyst directly to matters of principle that relate to what I have called 'preconditions of reference'. Recognition of the framework relativity of physical measurements and of claims made on their basis is today preeminently to be found among theoretical physicists whose work focuses on relativity and quantum theory; in both areas of study, framework relativity is not only recognized, but is firmly embedded in the very conceptual and mathematical structure of these theories.

Our interest in the following three chapters will be to reach a level and a form of understanding of relativity physics and quantum theory that is not generally developed in mathematical physics. The required mathematics commonly leads to the ability to represent the relevant physical phenomena in mathematical form and to manipulate the resulting formalism—in other words, to *calculate*, but not necessarily to be aware of or explicitly to analyze the referential preconditions which the mathematical representation and such calculations presuppose. The ability of the mathematical physicist competently to perform such calculations and the kind and level of understanding that are reached in the process comprise a fundamentally different kind and level than will concern us here. The ability to perform such calculations will not be our focus; as a result, the reader not trained in advanced mathematics should be neither intimidated nor disadvantaged. We shall, in part, be interested in describing how the mathematics functions in its referential capacity, not how to use it to perform those functions. In doing this, our specific object will be to identify as clearly as possible the preconditions of reference which the physical theories in question presuppose for their possibility and therefore for their meaning. Our focus, then, concerns the referential sub-structure, the metalogical scaffolding, of special and general relativity and of quantum theory.

In the next three chapters, we shall take note of a number of important theoretically fundamental connections and affinities that exist between, on the one hand, the metatheory formulated by the metalogic of reference and, on the other hand, the three physical theories, the special and general theories of relativity and quantum theory.

## 26

# Relativity Physics as Seen through the Lens of the Metalogic of Reference

### 26.1 Introductory comments

I preface the discussion that follows with several general comments and reminders of relevant results that have so far been reached:

First, it will not be my intent to give an exposition of the special and general theories of relativity, for a great number of explanatory works already exists, ranging from popular to highly technical; of these, many date back to Einstein's published work, and many more have of course appeared since.<sup>293</sup> An educated layperson's general understanding of relativity theory will be assumed. The discussion that follows supposes on the part of the reader—as elsewhere in this book—patience, tenacity, and a fairly high degree of abstract thinking, although, as noted, an advanced level of mathematical training will not be needed. Einstein (1921, p. 782) assured readers, as I will also do my best to accomplish, "... the whole ascent is composed of small, almost self-evident steps of thought."

Second, a few preliminary thoughts and reminders are in order to establish the context for the discussion that follows.

Readers will recall ( $\{5\}$ ) that in the present study "frames of reference" are understood to function as systems of coordination; such systems are not limited to those familiar in mathematics which have well-defined, quantitatively calibrated axes. We have come to understand that a range of possible objects of reference is determined by the referential capabilities of a given frame of reference; the concept of possibility, understood in these terms, is framework-relative in the sense that a frame of reference establishes a range of possible objects of reference; the preconditions of reference that define a given frame of reference establish what is possible from that standpoint.

---

<sup>293</sup> Among these many publications, for readers unfamiliar with relativity physics, and to keep this list short and limited to Einstein's own work, there are Einstein's widely read informal summary (Einstein, 1920/1917), his introductory brief overview (Einstein, 1921), his somewhat more technical book (Einstein, 1922), and his advanced exposition (1916a).

When the preconditions of reference that establish the basis for a given frame of reference are identified, what is possible in terms of that reference frame is made explicit ({7.3.10}).

We have also seen that for many reference frames it is possible to *transform* the way in which an object of reference is identified in a manner that conforms to the identification requirements of another system, or of other systems, of reference. A distinction was therefore drawn between compatible and incompatible reference frames: For two reference frames to be compatible, the means must exist whereby it is possible to translate the identity of an object given in one frame into a correspondingly identified object in the other frame, and such a transformation must be such that it preserves the identity of the object so that it is justifiable to claim that the two reference frames identify “the same object.” It was further noted that two frames of reference are *isomorphic* if and only if the two frames are so constituted that any object identified in the one framework can be placed in a one-to-one relation with an object in the other in a manner that satisfies criteria that qualify the two objects as being one and the same object. Two frames of reference that are not compatible in the above sense may nonetheless be considered *complementary* when, from the standpoint of a third, meta-framework, the set of objects identifiable in one frame qualifies as the same set of objects identifiable in the other ({10.5, 11.1}).

We shall find that the above results relate in specific but in different ways to the referential sub-structure of the special and general theories of relativity.

Before we proceed, we need to define the concept of identifying reference as it is represented in its more specialized physical variety. Up to this point, we have understood that identification/identifying reference obtains when that which is described is determined to be that to which there is reference. We recall that such identification has been broadly understood; it may range from unique to general, from vague to abstract, or it may consist in specification (again, precise or vague) according to rule ({8.6}). The identification of an object of reference may, in other words, be specific and unique to a particular object, or it may be subject to ambiguity, indefiniteness, or probability; no matter what its degree of specificity, the object’s identity is, in principle, a function of the frame of reference in terms of which it can be identified ({10}). Further, we have recognized that identifying reference to an object is only possible when such identification is fixed or defined within its presupposed framework in a way that assures the possibility of re-identification of the same object; such re-identifiability requires that a frame of reference permit temporally successive, reiterated reference ({22.1}). We have also come to see ({22.1}) that in order for any object of reference, physical or otherwise,

to be identified, it must also be re-identifiable, and that hence reference necessarily has a temporal basis. We have seen that any object of reference—that its very identity—is a function of the set of parametric constraints within which its identification can occur.<sup>294</sup> We shall now apply these results in a physical context in which physical events, objects, and related phenomena are in view. With these reminders of our abstract understanding of identifying reference, we turn now to the specific class of physical objects of reference studied by physics and in particular by relativity theory.

The objects of reference to which the special and general theories of relativity have been developed to refer are such that they necessarily involve physical processes of detection and measurement applicable to physical content, which may consist of material objects, events, energy levels, the properties of fields, etc. Methods of measurement are required in order to detect and to quantify physical phenomena; detection and quantification are of course the express purposes of physical measurement. Such detection and quantification comprise the basis for the identification of physical phenomena, and, as we have previously come to understand this in theoretically abstract terms, identification presupposes one or more appropriate frameworks of reference; the same will hold true for frameworks of reference that supply physical means of detection and quantification.

The conventional understanding in physics of a “frame of reference” (or “reference frame”) may for the purposes of this chapter be conceptualized as involving two distinguishable but usually interrelated aspects: There is of course the physical basis of such a frame of reference which provides both the physical means to determine, in relation to that reference frame, the spatial and temporal orientation of the set of physical objects to be identified, as well as explicit procedures in relation to which measurements of those objects can be standardized and hence communicated unambiguously among physicists. In addition, but on a purely conceptual level, we recognize the role of an open-ended multitude of alternative abstract, formal coordinate systems in terms of which it is possible to represent or designate the physical phenomena given in a specific physical reference frame. As did French physicist and mathematician Louis Marcel Brillouin (1854–1948), we may therefore draw a distinction between the wholly mathematical conception of coordinate systems and that of physical frames of reference.<sup>295</sup> The description of a physical reference frame may commonly involve, for example, the specification of a relationship between a set of physical phenomena and a detection/measuring apparatus or human observer; such a reference frame is called an ‘observa-

---

<sup>294</sup> Summarized in {13}.

<sup>295</sup> Brillouin (1970, p. 49).

tional reference frame'. The physical phenomena identifiable from the standpoint of such a physical reference frame may, however, be specified by means of any number of alternative formal coordinate systems. In this sense, the concepts of physical reference frame and of formal coordinate system may be distinguished. In the specialized context of physics, and specifically in that context, we may think of a coordinate system as a conceptual grid that is superimposed on or applied to quantitative measurement data derived from a physical reference frame.

## **26.2 The problem of correlating phenomena from the standpoint of different reference frames**

The special and general theories of relativity characterize, as we shall see, physical reference frames in terms both of their states of motion and their time measurements, while relationships among such physical reference frames are formulated by means of coordinate systems chosen because they are capable of *translating* physical references in one reference frame to those in one or more other physical frames. In particular, in order that physical laws may be formulated that apply equally to all physical reference frames, a physical phenomenon given in one reference frame must be correlated with what is judged to be a corresponding physical phenomenon given in other reference frames; in other words, there is a need, one that I've emphasized (*{10, 21, 22}*), to be able to ascertain that "one and the same" object of reference is involved. Here a conceptually important matter is involved, one often overlooked or dismissed without explicit study by both many theoretical physicists and philosophers of physics. How it is to be possible in the above sense to correlate different physical reference frames, which may be in different states of motion and possess different time keeping means, poses a central problem of *translation between frames of reference*: This is a fundamental problem of reference whose articulated solution is not to be found in Einstein's special and general theories of relativity. It is a problem that readers should bear in mind as they read the sections that follow.

One of the few physicists to have directed attention to this problem was P. W. Bridgman. The majority of physicists, on the contrary, tend to disregard the issue that is at stake, either passing over it in silence, or simply dismissing it as did physicist Abraham Pais, whose response to the problem is fairly typical. Pais referred to the "ability to correlate information obtained from two different reference frames" as an ability that is required in order to claim that "the same" physical phenomenon is identified by both reference frames. The only attention Pais gave to this requisite "ability" is via a footnote in which he

stated: "I trust that the term *the same* will not cause confusion" (Pais, 1982, p. 199n).

Bridgman's recognition of the significance of the problem stands in stark contrast. Since the conceptual issue is rarely discussed by physicists themselves, I quote him at some length:

The process of correlating the coordinates of a definite *happening* in one frame with those in another is what constitutes the transformation of coordinates. This correlation can be unequivocally made because it is one of my unanalyzables to be able to apprehend the coincidence of the mesh points of one framework with those of another....

It is obvious that with a setup like this no physical conclusions whatever about the happening can be drawn merely by passing from the coordinates in one framework to those in another. But what is usually done is somewhat different. Two observers are imagined, each with his own frame of reference. Each observer determines in his framework the coordinates of a succession of physical happenings, and then the transformation of coordinates is accomplished by correlating the coordinates obtained by the two observers for what they agree are the *same* happenings. That is, the unanalyzables of the two observers have to possess, in addition to the intuitively apprehendable properties for each observer already implied, the further property that two different observers can unequivocally agree on a "sameness."

What does the assumption of this possibility involve? It is not as simple as it might appear, because the things to which "sameness" is being ascribed are unanalyzables, and the intuitive operations which define the unanalyzables are certainly not the same, whatever that may mean, for the two observers. If one observer sees a flash of yellow light at a certain point of his framework, the other observer moving with high relative velocity may apprehend only a flash of infrared radiation by feeling a glow of heat on his finger. I think it would be difficult to persuade two observers to call an elementary event the same which was received by two totally different senses. It seems to me that the only basis for a secure judgment of "sameness" is a certain amount of discreteness in the elementary event. If happenings are discrete

enough, one observer can say to the other “Something just happened to me,” and the other can reply, “Something just happened to me too,” and by definition we can call this the same event, provided no confusion arise from the overlapping of other events. But other events always do overlap to a certain extent, and it would probably be very difficult to give a rigorous definition of what would constitute prohibitive overlapping, or exactly what “enough” discreteness means....

These considerations I believe may well raise doubts as to the fundamental assumption. We may be willing to assume that a given physical situation may be analyzed into discrete unresolvable elements by a single observer with sufficient lack of ambiguity for the purposes of certain physical theories, and at the same time question whether two different observers can analyze the same situation into discrete events in such a way that the two observers can attach a property of “sameness” to the events in addition to the other properties. Operationally it is evident that much more is involved in the second situation than in the first. It seems to me that the existence of this property of “sameness” acquires its plausibility in terms of a feeling for an underlying “reality,” which seems to me almost metaphysical in character, and to which I can see no way of giving sufficient operational sharpness. (Bridgman, 1936, pp. 75-79)

This passage was written by a physicist known for having demanded clarity concerning the fundamental concepts upon which physics relies. And yet we see him speak of the “unanalyzables” and “intuitively apprehendable properties” that purportedly are relied upon when individual observers conclude that they have “the same” physical phenomenon under consideration. These strikingly unclear and muddled expressions were—and this *is* clear—intentionally contrived by Bridgman in order to bring out the fact that the recognition of “sameness” is not nearly as precise or as well understood as it should be. And there is no doubt that better understanding is needed, given that the recognition of “sameness” is basic to correlations that we shall, in discussing relativity physics as well as quantum theory, need to make with respect to different observational reference frames. Bridgman’s implication is that a surreptitious metaphysical step is involved that attributes an “underlying reality” to physical phenomena when they are identified by means of different reference frames, and when they are then judged to comprise “one and the same object.”



In previous chapters, I have alluded to this problem, which we should now consider more explicitly, albeit only briefly. Reference to “the same object” has so far concerned us in two distinguishable contexts: On the one hand, we’ve seen that identifying reference to any object is only possible when the object is in some way fixed or defined within a presupposed frame of reference which makes it possible to re-identify “the same object.” On the other hand, we’ve come to understand “compatible reference frames” as frames of reference in terms of which it is possible to refer to “the same object.” These two uses of the phrase ‘the same object’ are of course different; it is the second of these that relates to the problem of correlating phenomena that are given from the standpoint of different reference frames.<sup>296</sup>

We recall ({10.5}) that isomorphic frameworks of reference make it possible to place an object identified in one framework in a one-to-one relation with an object in the other framework in a manner that satisfies criteria—or correlation procedures—that enable us to claim that the two objects comprise “the same object.” If we are to find a solution to the problem of correlating phenomena identified from the standpoint of different reference frames, that solution must evidently be found in looking more closely at these criteria. For the purposes of earlier chapters in this study it was enough to say (as in {21.1.1}) that these criteria or correlation procedures express certain abilities, usually exercised by us implicitly, in a way that has become habitual, usually automatic, and second-nature: We give little or no thought when we correlate objects identified by means of sight, touch, hearing, or another sense, and then, “when circumstances seem right,” we claim that these variously given objects relate to “the same object.”

But we now face the need to understand both the correlation criteria and those additional “circumstances that must be right” before the criteria can be applied. It would lead us too far astray in the present study to analyze the complexities of this issue; here, I refer readers to a separate monograph (Bartlett, 2015). The conclusion reached in that work is that recognition of similarities among objects of reference—reference to “the same object,” constituting a maximal degree of similarity—is inescapably ambiguous and requires recourse to selection procedures whose responsibility it is to disengage individual objects of reference from the complex, interconnected background

---

<sup>296</sup> The first sense relates to the understanding (see definition D 2, {11.4}) that identifying reference establishes that what is ascribed (one or more properties, relations, a description, etc.) and that to which ascription is made are “one and the same” (i.e., identification): In such a case of identification, the content of what is ascribed and the subject of the ascription are seen to be identical; only a single framework of reference is involved and no correlation between distinct reference frames is needed.

of a multiplicity of other objects of possible reference, and by doing this, to eliminate what Bridgman expressed as the “confusion” that can “arise from the overlapping of other events.” If I am right that “inescapable ambiguity” is necessarily involved, then the criteria, the correlation procedures, that are at issue here can only be understood as essentially heuristic, not subject to algorithmic formalization. If so, then the problem of correlating phenomena given in different reference frames is not, as Bridgman seems to have implied, “metaphysical” in nature, but instead points to pattern recognition abilities that must be presupposed, whether on the part of human or non-human observers.

As we proceed, we should keep this matter in mind in the context of the theoretical physics we shall discuss: When correlating physical phenomena identified from the standpoint of different observational frames of reference, we appear to be forced to rely upon non-formalizable abilities<sup>297</sup> on the part of human or non-human observers in order to recognize commonalities among those phenomena, commonalities that are sometimes sufficient so that we may judge that they pertain to “the same object” from the standpoint of compatible reference frames. These general and abstract considerations will apply directly and on a fundamental level as we consider the referential preconditions established by the special and general theories of relativity.

### 26.3 Fundamental properties of the special and general theories

In several ways Einstein’s two theories of relativity are unusual among theories that have been developed in physics: To begin with, they were developed by him during a period marked by his rare degree of isolation from experimental work. Both the special and the general theory resulted almost entirely from purely theoretical reflection, analysis, and mathematical formalization. In retrospect, as Einstein expressed this: “It is my conviction that *pure mathematical construction* enables us to discover the concepts and the laws connecting [physical events], which give us the key to the understanding of the phenomena of Nature” (1934/1933, p. 167, italics added). After years of intense labor during which he sought to extend the special theory of relativity

---

<sup>297</sup> Such abilities, as shown in Bartlett (2015), are “non-formalizable” in the sense that pattern recognition resulting from their use is not deductively derivable or obtainable by strictly rule-driven, decision-effective algorithmic means within any formalized system. This, however, does not of course imply that the abilities that comprise pattern recognition both of similarities and of the equivalence of objects of reference cannot be understood as a consequence of the analysis of recorded large sets of data, something which is successfully realized by artificial intelligence software designed for this purpose.

to the general theory, he commented that he found himself “in a position *to deduce, in a purely theoretical way*, the properties of the field of gravitation” (Einstein, 1920/1917b, p. 50, italics added, author’s translation). Describing Einstein’s achievement in relativity physics, H. A. Lorentz wrote that Einstein’s work possesses a “technical simplicity in the critical assumptions which makes the wealth of deductions astonishing. *It is a case of an advance arrived at by pure theory....* Einstein arrived at his theory through a train of thought of great originality....” (Lorentz, 1920/1919, pp. 24, 30, italics added). Lorentz went on to remark, in a section of his paper entitled “Difficulty Exaggerated,” that “the basic ideas of the theory [of relativity] are really clear and simple” (p. 59).

“Clear and simple deductions made on the level of pure theory,” “composed of small, almost self-evident steps of thought...”—such observations as these about Einstein’s revolutionary work are not often encountered in commentaries, which instead tend more often to promote the view that his special and general theories of relativity are intellectually and mathematically intimidating, that they are complex and difficult to comprehend, and that they lead to highly counterintuitive results (which for many readers these results can be, much like certain of the results reached by the metalogic of reference). As one studies the development of Einstein’s thought and comes to recognize this combination of clarity, simplicity, and incremental assembly from nearly self-evident steps, one cannot at times but wonder to what extent the inhabitants of an impenetrable cloud-shrouded planet might similarly be able to deduce, through purely theoretical means, conclusions about the physical universe not accessible by means of their observations.

I wish to emphasize this purely theoretical basis of the physics of relativity for several reasons:

H. A. Lorentz, in his concise account of Einstein’s work, wrote:

It is certainly remarkable that these relativity concepts [the denial of absolute motion and absolute simultaneity] ... have found such a rapid acceptance.

*The acceptance of these concepts belongs mainly to epistemology.* (Lorentz, 1913, p. 23, italics added)

Epistemology is, of course, a purely theoretical area of investigation, and Einstein’s approach in attempting to solve what we shall come to understand as the central referential problems of relativity physics may, in great measure, as we shall see, be considered to be epistemological in nature. Einstein was seriously interested in epistemology, and he is known to have devoted a signi-

ficant amount of time to its study.<sup>298</sup> Toward the end of his life, Einstein wrote: “Science without epistemology is—in so far as it is thinkable at all—primitive and muddled” (Einstein, 1959/1949a, p. 684).

Despite his evident concern and respect for epistemology, Einstein’s occasional appraisal of general philosophy was far from glowing: “Is not all of philosophy as if written in honey? It looks wonderful when one contemplates it, but when one looks again it is all gone. Only mush remains.”<sup>299</sup> Perhaps he may have judged epistemology to comprise a field of investigation that can—or perhaps should—be separated from general philosophy and then developed independently in much the same way as physics became autonomous of natural philosophy a few centuries ago, or perhaps in his later years his evaluation of philosophy soured and his positive regard for philosophy diminished; we do not know.

But, in any event, Einstein’s approach in developing the special and general theories of relativity was, as Lorentz and others have noted, markedly epistemological, with an explicitly stated primary concern for “first principles” and logically compelling deductions made from them. Such an objective has of course also characterized large portions of traditional philosophy. As discussed in the first chapter of this book, it has been a classical philosophical goal to attain a level of conceptual universality and generality through a study of the most fundamental concepts, the first principles, the premises of human thought and knowledge, principles that are the necessary underpinnings of rationality and knowledge. Einstein’s aim within physics accorded with this classical ideal; he, too, wished to develop an all-encompassing general theory that would identify, based on a minimally small set of first principles, laws of nature which both would be logically derivable from that comprehensive theory and would hold irrespective of the particular conditions of individual physical frames of reference. He sought to find such invariants that hold true no matter what the individualizing conditions of physical reference frames might be.

In {10}, I mentioned Einstein’s belated regret that he had chosen the name ‘relativity theory’ instead of ‘invariance theory’ (*Invariantentheorie*).

---

<sup>298</sup> In school, Einstein studied Kant’s work. Later in Bern he met regularly c. 1902 with fellow members of what they called the ‘Akademie Olympia’ to read and discuss Spinoza, Hume, Kant, Mill, Mach, Poincaré, Avenarius, and others. (See Pais, 1982, pp. 47, 133, 318-319.) In the 1920s, Einstein wrote a number of reviews of books dealing with epistemology by Steven Weinberg and Josef Winternitz, which make evident his knowledge of Kant; he also wrote several introductions for philosophical works. He studied Bertrand Russell’s theory of knowledge, and reviewed Meyerson’s (1925) *La déduction relativiste* (its English translation (Meyerson, 1985/1925) includes that review).

<sup>299</sup> Quoted in Rosenthal-Schneider (1980, p. 62).

Bertrand Russell commented that the naming of relativity theory may have been inappropriate:

Perhaps the name [the ‘special and general theories of relativity’] is unfortunate; certainly it has led philosophers and uneducated people into confusions. They imagine that the new theory proves *everything* in the physical world to be relative, whereas, on the contrary, it is *wholly concerned to exclude what is relative* and arrive at a statement of physical laws that shall *in no way depend upon the circumstances of the observer*. (Russell, 1958, p. 16, italics added)

Many years before Russell expressed this criticism—and even before Einstein’s own admission of regret over the naming of his theories—mathematician Adolf Kneser (1918) wrote that it was Einstein’s (as well as Lorentz’s) achievement to have directed

... attention toward the things that must *not* be considered from the relative point of view [*die sich der Relativierung entziehen*—i.e., which *resist* or *evade* relativizing]. The principle of relativity is, as a matter of fact, the principle of *the nonrelativity of the real*; it demands that the reality implied by the observed phenomena of nature remain *immutable* with respect to possible modifications of viewpoint and system of measurement.... (quoted and translated in Meyerson, 1985/1925, p. 50, italics added)

Due to the frequent confusion and ease of interchange of the notions of “relativity” and “relativism” (especially at a time now dominated by a steadfast Protagorean rejection of absolute truth), Einstein’s decision to call his theories ‘of relativity’ has led many to miss the essential fact that he was concerned with principles that are not at all relative, but instead are genuinely *absolute*: Far from endorsing—indirectly or even implicitly—any form of relativism, Einstein’s major contribution to physics seeks to identify *absolute* physical laws that are universal and covariant with respect to the variable physical conditions of arbitrarily chosen reference frames. (We shall discuss the concept of covariance later in this chapter.) The primary concern and focus that defined and absorbed Einstein’s attention related to a search precisely for such *invariants*. The purpose of physics, as Einstein conceived it, is to explore and describe “this world of *absolute invariants*” (in the words of Meyerson,

1985/1925, p. 51, italics added).

It should be evident that in various ways it has been a regrettable and misleading choice, inconsistent with Einstein's own intent, to call the result of his search 'relativity theory'. Its name is now indelibly etched in history, but it is important for us to recognize how that name masks the central purpose of relativity physics.

Relativity physics, in this way, bears a significant affinity with the metalogic of reference: Much the same fundamental motivation underlies the metalogic of reference as a theory of formal ontology ({8.6}), that is, to describe *invariant principles* that govern the identifiability of objects, invariants that can be *transformed without loss of validity* from one reference frame to another, irrespective of the nature of those objects. The invariants that concern us in this study are *invariants that govern the framework relativity of reference in all reference frames*, much as the invariants of Einstein's physics of relativity describe *invariant principles whose validity and form remains the same in all physical reference frames*.

In the sections that follow, we shall come to see how Einstein's progression from the special to the general theory went hand-in-hand with increases in the sophistication of his referential analyses, and how his analyses, in particular instances, parallel those we have made in this study and lead to results in physics that, perhaps surprisingly, agree with certain of this study's conclusions, conclusions that come about from the application of an essentially different, purely theoretical, and non-physical method.

#### 26.4 Steps from Newtonian physics to general relativity

I have suggested that Einstein's approach in formulating invariant physical principles was fundamentally epistemological. He described the objectives of physics in the following philosophically abstract terms:

[Physics] seeks ... to build up a logical system, based on as few premises as possible, which contains all laws of nature as logical consequences. This system, or rather the structures occurring in this system, is *coordinate with [zugeordnet]* the objects of experience. (Einstein, 1928, p. 252, italics added)

Even such a two-sentence description of his conception of physics clearly has pure theory at its core: Einstein's objective was not to develop a general physical theory inductively inferred from empirical data, but rather the reverse, to formulate a general physical theory both which comprises a logical

system from which physical laws can be deductively derived, and which functions as a *system of coordination* with physical experience. The special and general theories of relativity accomplish this by means of carefully designed steps of coordinate transformation, as we shall see.

In {5.1}, I briefly discussed Reichenbach's conception of knowledge as coordination, a conception which he proposed in his short, not fully developed, epistemological analysis of Einstein's relativity theory. We recall that Reichenbach considered physical objects to comprise "reference structures" that function in essentially coordinative terms. His notion of physical objects as "reference structures" remained an incomplete notion that perhaps sought to point in the direction of what we've come to describe as the framework relativity of objects of reference.

An epistemologically analytical understanding of the fundamental coordinative function of systems of physical reference in relation to physical phenomena was central to Einstein's ability to progress beyond Newtonian physics. Let us see, in the terms with which we have become familiar in this study, how an epistemologically reflective analysis led Einstein first to his special theory, and then to his general theory of relativity.

We've noted that identifying reference to physical phenomena—i.e., physical reference to such phenomena in a manner so that they may be identified and re-identified—requires specification provided by measurement. It has been the heritage of classical physics, a heritage extending back to antiquity, to require that such specification involve measurements of *distances* and of *times*. It seemed evident, intuitive, and unavoidable that, in order to refer identifyingly to a physical phenomenon, that phenomenon must somehow be situated in terms of its spatial and temporal dimensions, and for this to be possible, a system of reference was needed that supplies the means for such specification—in other words, a physical reference frame from the standpoint of which relevant spatial and temporal measurements can be made.

The central role of spatial and temporal measurement acquired the character of inescapable obviousness, so much so that careful, analytical attention was seldom paid to certain of the most elementary assumptions upon which physical reference by means of measurement relies. It was Einstein's particular talent, justifiably applauded as genius, to have questioned and analyzed how certain of the most fundamental assumptions pertaining to the identification of physical phenomena are employed and built upon by physics. As we shall see, when questioned carefully, in Einstein's recognizably epistemological approach, certain of the central, conventional, and purportedly self-evident assumptions of Newtonian physics lost their persuasiveness and were seen to

be unjustified, in this way leading first to the special theory and eventually to the general theory of relativity.

### 26.5 The starting point: The Galilean transformation

We begin with a short reminder of the historical context out of which Einstein's work evolved. The traditional fundamental status accorded to the dimensions of space and time encouraged Newton to accept the unquestioned belief that there exists an absolute space and an absolute framework of time in relation to which all physical bodies move and events take place. Like most physicists, he wished to formulate physical laws that would serve as universal principles valid for all observers. The laws he formulated, which he believed in this sense to comprise invariants, presupposed such a framework consisting of a single, absolute, all-embracing space and a single, absolute, all-encompassing time in terms of which physical phenomena can be situated and thereby identified. As a result, when—using today's terminology—two observational reference frames seek to refer to what is judged to be the same physical object or event,  $P$ , Newton assumed that a simple and direct *translation*, or *transformation*, can be made that would correlate the perspectives of the two observational reference frames. If, for example, two observational reference frames  $S_1$  and  $S_2$  are moving apart with respect to one another in simple unaccelerated relative motion in a straight line with a velocity  $v$  ( $S_1$  and  $S_2$  are known as 'inertial reference frames'<sup>300</sup>), then in Newtonian mechanics a simple transformation can be made that accomplishes this correlation of different perspectives: From the standpoint of  $S_1$  let us say that the physical object or event  $P$  can be specified by coordinates  $(x_1, y_1, z_1, t_1)$ , and from the standpoint of  $S_2$ , by coordinates  $(x_2, y_2, z_2, t_2)$ . For simplicity, we assume that the direction of movement of  $S_1$  and  $S_2$  in relation to one another is only along the  $x$ -axis, and that  $t_1 = t_2$ . Then the spatial position of  $P$  would be specified simply by  $x_1 = x_2 - vt$ . Relative to either of these two inertial frames of reference, according to the classical mechanics of Newton, physical phenomena will conform to the same general laws of physics as they do with respect to the other reference frame. Einstein referred to this invariance as "the principle of relativity (in the restricted sense)" (Einstein, 1920/1917a, p. 13); it expressed the recognition of a limited level of invariance that preceded the special theory of relativity.

---

<sup>300</sup> Observational frames of reference are of two kinds: *inertial*, at rest or moving with a constant speed in a straight line, and *non-inertial*, experiencing acceleration or moving in a curved path (or, equivalently, in a gravitational field). An inertial reference frame is also called a 'Galilean reference frame'.



This transformation has come to be known as the Galilean transformation, named by physicist Philipp Frank in 1909.<sup>301</sup> The Galilean transformation provided accurate results in Newton's physics, but as Einstein came to see, the transformation applied only to physical reference frames that move at considerably less than the speed of light.<sup>302</sup> The Galilean transformation provided a way to translate between coordinate systems employed by different observational reference frames in relative uniform motion with respect to one another. But it was a transformation whose referential basis Newton failed to examine adequately. The revolution in physics for which Einstein was responsible led from a model based on intuition and unexamined assumptions to one based on rigorous analytical reasoning, reasoning which in fact centered on an analysis of preconditions of physical reference.

### 26.6 The first step: The Lorentz transformation

Newtonian physics is characterized by its need to attribute to mass a space and a time that are thought to possess real independent existence.

– Albert Einstein (1965/1956, p. 155)

We have noted that the central motivation behind Einstein's "theory of invariance"—his two theories of relativity—was his wish to formulate physical laws in a manner that is invariant, valid no matter what the individualizing conditions of an observational reference frame might be. He distinguished two stages in his search for such invariant principles, initially the special and, some years later, the general theory.

The special theory more informatively could be called the '*restricted theory*' since it sought to formulate laws of nature that remain invariant with respect to observational reference frames whose motion is *restricted* to uniform change of position;<sup>303</sup> the general theory of relativity, in contrast, lifted this restriction and formulated laws of nature that are invariant irrespective of the states of motion of observational reference frames.

The restricted theory of relativity comprised a significant advance beyond classical mechanics, whose objective Einstein characterized as describing "how bodies change *position with time*" (Einstein (1965/1956, p. 15, italics

---

<sup>301</sup> Pais (1982, p. 140n).

<sup>302</sup> Of course, not necessarily limited to light in the visible spectrum of electromagnetic signals.

<sup>303</sup> The French have made this choice and have generally translated Einstein's German "*spezielle Relativitätstheorie*" as "*théorie de la relativité restreinte*."

added). As we shall see, the concept of position (implicating the concepts of length and distance) and that of time were to play essential and fundamental roles in the restricted theory of relativity.

Einstein recognized the need to question and then to reject the two fundamental assumptions of pre-relativity Newtonian physics: that time is absolute (that the time-order of events from the standpoint of a particular frame of reference is the same time-order from the standpoint of any other frame of reference) and that length is absolute (that a length measured relative to one frame of reference will have the same length relative to any other frame of reference). These essentially referential physical assumptions had maintained a largely unselfconscious rigidifying grip on the minds of physicists since Newton, and stood in the way of the further development of physics. They were tenuous assumptions largely because they lacked clear definitions formulated in measurable terms.

In demanding measurement-based definitions of length and time, Einstein might give the impression that he insisted upon an operational approach, but this was not the case. In his response to operationalist P. W. Bridgman's criticisms, Einstein wrote:

In order to be able to consider a logical system as physical theory it is not necessary to demand that all of its assertions can be independently interpreted and "tested" "operationally"; *de facto* this has never yet been achieved by any theory and can not at all be achieved. In order to be able to consider a theory as a *physical* theory it is only necessary that it implies empirically testable assertions in general. (Einstein, 1959/1949a, p. 679)

What distinguishes length and time as notions specifically in need of operational, measurement-based definitions is the fact that, in the absence of such definitions, physical identifying reference to material bodies, events, energy levels, properties of fields, etc., is rendered impossible. As we've already noted, for physical objects of reference to be identifiable metalogically requires physical processes of detection and measurement; such detection and measurement procedures comprise the necessary basis without which the identification and re-identification of physical phenomena are impossible. We note that the rejection of the preceding statement is manifestly projective: For instance, the following claim is metalogically projective: "Event *e* has a physical identity that is such that *e* occurred at a place and time, neither of which can in principle be identified." Putative physical reference to event *e*

under such conditions is projectively meaningless. As we shall see, Einstein reached much the same conclusion, but by different means.

Given Einstein's objective to formulate empirically valid principles of nature that are invariant from the standpoint of all possible observational reference frames, there is a fundamental and unavoidable need to be able to *correlate* physical data acquired from the standpoint of different physical reference frames, and to do this in a manner that is independent of the states of motion or time-keeping measurements of those frames of reference. To render such correlations possible, the notions of time and length stood in need of definitions linked with actual means of measurement. The role of the *definition* of these notions in this context determines the very basis for physical reference, as well as the physical consequences that follow from these conceptually basic definitions.

Since the referential ability to correlate—to coordinate—distinct physical reference frames that may be subject to different states of motion defines the problem to be solved, Einstein faced the need to formulate operational definitions of length and of time that make such correlations possible. To be able to judge that “the same” physical object, event, energy level, property of field, etc., is identified with respect to two different observational reference frames, there is an unavoidable need to be able to specify times and distances, the spatial and temporal coordinates, of that “same” physical object of reference. In connection with the dimension of time, Einstein reasoned:

In order to measure time, we ... suppose a clock,  $U$ , present somewhere, at rest relatively to [a reference frame]  $K$ . But we cannot fix the time, by means of this clock, of an event whose distance from the clock is not negligible; for there are no “instantaneous signals” that we can use in order to compare the time of the event with that of the clock. In order to complete the definition of time we may employ the principle of the constancy of the velocity of light in a vacuum. (Einstein, 1922, p. 28)

Recourse to signals that travel at the speed of light was certainly not a requirement derivable from purely logical considerations; in fact, Einstein frequently made it clear that his reliance upon the constancy of light signals in a vacuum in order to establish spatial and temporal correlations between different physical reference frames was a “*supposition*.” He wrote, for example:

[Special relativity] accepts from the theory of Maxwell-Lorentz the *supposition* of the constancy of the velocity of light in a vacuum. In order to place this *supposition* in accord with the equivalence of inertial systems (the principle of special relativity), it is necessary to abandon the absolute character of simultaneity. (Einstein, 1965/1956, p. 169, italics added, author's translation)<sup>304</sup>

It is important to emphasize this fact, that the role of the constancy of the speed of light in a vacuum, in relation to inertial systems of reference, is precisely that of a supposition or postulate. Although independent experimental evidence exists to support this claim,<sup>305</sup> strictly from the standpoint of pure theory, Einstein was intellectually honest in recognizing that, relative to inertial reference systems, the constancy of the speed of light in a vacuum had for him the status of an assumption.

Once this postulate is made in tandem with the postulate of the principle of relativity understood in the restricted sense (again, that the laws of nature remain the same in any inertial frame of reference), several consequences follow which will be familiar to readers already acquainted with special relativity. They include the phenomena of time dilation and length contraction, and the relativity of simultaneity, all of which become detectable at high velocities, and are appreciable at velocities that are a meaningful fraction of the speed of light. Once it is assumed that the speed of light *in vacuo* remains a fixed constant in relation to all inertial reference frames, and it is accepted that light signals are to be relied upon in coordinating events from the stand-

---

<sup>304</sup> A related example may be found in Einstein (1920/1917, p. 42) where he labeled a "*postulate*" the claim that "the velocity of transmission of light *in vacuo* has to be considered equal to a constant *c*."

<sup>305</sup> Classical examples include the negative result of the famous Michelson-Morley interferometry experiment of 1886, later repeated in a variety of more sophisticated ways and confirmed by others, none of which experiments found evidence that the speed of light is affected by the Earth's orbital speed; classical examples also include astronomer Willem de Sitter's 1913 research relating to double stars, which supported the claim that the speed of light does not depend on that of the source and is constant in (comparatively) empty space.

By 1915, Einstein came to conclude that the speed of light is constant in a vacuum under gravity-free conditions. In the presence of gravitational fields, however, it is important to realize that the speed of light is variable. The variability of the speed of light in a gravitational field is currently employed, e.g., in laser gyroscopes of inertial navigation systems; the phenomenon is known as the Sagnac Effect.

Einstein's "supposition" or "postulate" continued to attract research interest among physicists during the second half of the twentieth century; some of them have proposed alternative theories of the variable speed of light, e.g., physicists Robert Dicke, Jean-Pierre Petit, Andreas Albrecht, João Magueijo, and others. For more detailed discussion, see Koks (2006, Chap. 7).

point of different reference frames, then if two physical reference systems are in rapid motion with respect to one another, events judged to be simultaneous in one reference frame will not be judged simultaneous by an observer in the other, and vice-versa. Time intervals are similarly affected as a function of the speed of movement of inertial reference systems in relation to one another: An observer considered to be in rapid movement in relation to another observer will find that his, her, or its measurement of an interval (the interval, e.g., between “ticks” of a time-keeping device) will be longer than the corresponding interval measured by the other observer, using an identical time-keeping device (and if this other observer is considered to be the one moving rapidly, the same result will be found, i.e., intervals between “ticks” in the moving system will be longer than those in the other reference frame). The same kind of result affects length measurements: From the point of view of a moving observer, lengths (also areas and volumes) will be measured to be contracted in the direction of motion in relation to a relatively non-moving observer. In addition to these theoretically derived consequences, all of these results of the special theory have also received strong experimental confirmation.

However, these physical results that follow from the special theory will not be our concern here, but rather the referential means that have made it possible to reach those consequences. Those “referential means” were supplied largely by H. A. Lorentz, and were used as a stepping stone by Einstein to formulate his special theory of relativity.

As we’ve seen, in accordance with the Galilean transformation, the correlation of the relative velocities of two reference systems is additive. To give an example: If from the standpoint of an inertial frame  $S_1$  an object  $o_2$  is moving with a velocity  $v_1$  in relation to an object  $o_1$  which is at rest relative to  $S_1$ , and if from the standpoint of a second inertial frame  $S_2$  which is at rest in relation to  $o_2$ , a third object  $o_3$  moves in the same direction as  $o_2$  with a velocity  $v_2$  away from  $o_2$ , then from the standpoint of  $S_1$ , the velocity of  $o_3$  is simply additive:  $v_1 + v_2$ .

However, when both of Einstein’s postulates are strictly adhered to, the additive result of the Galilean transformation no longer holds. Lorentz recognized that if light signals are relied upon to establish spatial and temporal correlations of events in inertial systems moving in relation to one another, and if those signals’ speed is stipulated to be a single unvarying constant irrespective of the motion of the inertial frames, a direct Galilean translation between the reference frames cannot be made, but instead the Galilean transformation must be replaced by a method of translation between the moving inertial frames that *incorporates* and *takes into account* the invariant

speed of light in a vacuum relative to inertial systems. The transformation that accomplishes this was named after Lorentz.<sup>306</sup> The derivation of the Lorentz transformation will not concern us;<sup>307</sup> instead, what is important in the context of the present study is the fact that now *a specific means for identifying physical events* has been specified—correlation of inertial systems through the use of light signals—and a limit that applies to that means of identification has been established as a required postulation—the finite, constant velocity of light in a vacuum.

A number of reference-based consequences follow when we make the constancy of the speed of light (always in a vacuum) a precondition, from the standpoint of inertial reference systems, of the identification of physical objects or events. Here are some of these consequences:

- Measurements of distances and times do not provide information concerning the measured properties of objects or events that is framework-independent, but instead such measurements supply information about relations of those objects or events to the physical reference frames from the standpoint of which those measurements can be made.
- The duration of an event or the length of an object has *no meaning* independently of the physical reference frames that are involved in their measurement; the length of an interval, the time-order of events, and the spatial length of an object are functions of the state of motion of those physical reference frame(s).
- The simultaneity of events has *no meaning* independently of the state of motion of the system of reference that is used. And, in general, “a time specification has *no meaning* unless it refers to a frame of refer-

---

<sup>306</sup> Below is a common formulation of the Lorentz transformation for two reference frames from the respective standpoints of which an event has coordinates  $(x, y, z, t)$  and  $(x', y', z', t')$ , with the velocity of the two reference frames here assumed for simplicity to be limited to the  $x$ -direction. The first frame observes the second as moving with the velocity  $v$  along the  $x$ -axis,  $c$  is the speed of light:

$$\begin{aligned}x' &= \gamma (x - vt) \\y' &= y \\z' &= z \\t' &= \gamma (t - vx/c^2)\end{aligned}$$

The so-called ‘Lorentz factor’ is  $\gamma = 1/\sqrt{1 - v^2/c^2}$ ; this factor makes clear that when  $v$  is much less than  $c$ , the factor plays an unimportant role and approaches the Galilean transformation described in the text above, but as  $v$  draws closer to  $c$ , its effect becomes significant.

<sup>307</sup> Einstein provided a simplified derivation in Einstein (1920/1917, Appendix I, pp. 115-120).

ence to which it is related” (Einstein, 1920/1917a, p. 18, italics added, author’s translation).

As Einstein commented on the latter claim:

We ... require a definition of simultaneity such that this definition supplies us with the method by means of which ... [the physicist] can decide by experiment whether or not [two events] occurred simultaneously. As long as this requirement is not satisfied, I allow myself to be *deceived* as a physicist (and of course the same applies if I am not a physicist), when I imagine that I am able to attach a *meaning* to the statement of simultaneity. (I would ask the reader not to proceed farther until he is fully convinced on this point.) (Einstein, 1920/1917a, p. 22, italics added)

It is important to emphasize that in passages like those quoted above, the consequences of the special theory repeatedly assert the *meaninglessness* of central Newtonian concepts. These assertions of meaninglessness are physically based; as we shall see later in this chapter, equivalent results are obtained by the metalogic of reference, but they are obtained in a non-physical manner, which does not rely upon empirical confirmation, and in a manner that cannot *not* be accepted without self-undermining referential inconsistency. In much the same way, we shall see later that Einstein’s general theory adds additional limitative but invariant consequences which entail their own relevant referential preconditions.

### 26.7 The second step: The shift to intrinsic identification

Thus far we’ve seen that Einstein’s special theory of relativity has a restricted application to physical reference frames that are inertial, that is, those not subject to acceleration or gravity. To extend his “theory of invariants” to include non-inertial systems was the purpose of the general theory—i.e., to find a comprehensive way to meet the requirement that all physical reference frames, no matter what their states of motion may be, should be equivalent in terms of the formulation of the laws of nature. He called the statement of this requirement the ‘general principle of relativity’.

In the first part of this chapter in which our focus has been the identification of physical objects and events by means of the special theory of relativity, we moved from the Galilean to the Lorentz transformation in order to

incorporate the postulated requirement of the constancy of the speed of light in a vacuum. During the years following his special theory, Einstein focused his attention on the indistinguishability, from the standpoint of the observer, of the observer's acceleration, on the one hand, and the observer's presence in a gravitational field, on the other. This indistinguishability is known as the equivalence of inertial and gravitational mass, which Einstein named the 'principle of equivalence'. Einstein's special theory's contraction of lengths and reduced rates of time-keeping devices in moving reference frames required modification once he recognized that these consequences become variable when acceleration or a gravitational field is involved. To be more specific, it was not hard for him to see that an object observed to be moving in a uniform straight line from the standpoint of an inertial system of reference is instead observed to be moving in a curved path from the standpoint of an accelerated reference frame. Similarly, when observed from the standpoint of an inertial system of reference, light signals in an accelerated physical reference frame, or in a gravitational field, generally do not follow straight lines as they do from the standpoint of inertial, Galilean frames of reference. A trajectory of an object is, in our terms, framework-relative—in fact, a “moving” object does not define any absolute, single, unique trajectory; trajectories are functionally determined by the reference systems in terms of which they are identified: “there is no such thing as *an independently existing trajectory* ..., but only a trajectory relative to a particular body of reference” (Einstein, 1920/1917a, p. 10, italics added). Nor is there any such thing as an absolutely “moving” object; motion itself is framework-relative.

Einstein's recognition of the principle of equivalence forced him to find a different way of correlating objects or events observed from the standpoint of systems of reference that are accelerated or are subject to gravitational fields. The reader will immediately see that the Galilean transformation, limited as it is to rectilinear motion in relation to inertial reference frames, could no longer be used when variable curved space-time paths are involved, nor would rectilinear coordinate systems continue to be useful. There was a need for a different way to correlate the identification of accelerated physical phenomena and, equivalently, those in a gravitational field. The way which Einstein found makes use of what I shall call the '*shift to intrinsic identification*'. It provides, as we shall see, a means of identifying physical phenomena that bears a very strong affinity with and resemblance to what in this study we've called 'referential fields'. The shift to intrinsic identification is therefore of particular interest to us in the context of this study.

The extension of the special, or restricted, theory to the general, or unrestricted, theory was conceptually challenging for Einstein, as it can be for



many readers today. Einstein commented, "...it was soon found impossible to link up the science of gravitation with the special theory of relativity *in a natural manner*" (Einstein, 1921, p. 783, italics added). Compared with the passage from Newton's physics to special relativity, the transition from the special to the general theory was unquestionably less direct and less easily won. Where the transition from Newtonian mechanics to special relativity can be directly understood in terms of Einstein's application of the Lorentz transformation, the transition from the special to the general theory required mathematics—which we shall come to understand as tools of referential analysis—with which Einstein was initially not familiar. The mathematical means were supplied principally by the combined application of contributions by mathematicians Carl Friedrich Gauss, Georg Friedrich Bernhard Riemann, Gregorio Ricci-Curbastro, and Tullio Levi-Civita.

The problem Einstein set for himself we've already described: to develop a comprehensive theory of the physical invariants of nature, invariants that apply equally to all physical reference frames, irrespective of their states of motion—and recognizing the equivalence of inertial and gravitational mass—therefore also irrespective of their presence in gravitational fields or their acceleration. To realize this goal would be to formulate a general theory of relativity.

This problem remained insoluble to me until 1912, when I suddenly realized that Gauss's theory of surfaces holds the key for unlocking this mystery. I realized that Gauss's surface coordinates had a profound significance. However, I did not know at that time that Riemann had studied the foundations of geometry in an even more profound way. (Einstein, quoted in Pais, 1982, pp. 211-212).

Einstein realized that the curvilinear paths traced by light signals passing through a gravitational field required a departure from Euclidean geometry. In the presence of gravity, it is not possible to measure length by means of "rigid bodies" that conform to Euclidean geometry; physical means of measurement by means of rigid rods, useful in inertial systems, are influenced (lengths are distorted) in the presence of a gravitational field. As Einstein expressed this: "In gravitational fields there are no such things as rigid bodies with Euclidean properties; thus the fictitious rigid body of reference is of no avail in the general theory of relativity" (Einstein, 1920/1917a, p. 98). The measurement of time is furthermore similarly affected.

In this chapter, we've understood coordinate systems as they are

employed in physics as abstract conceptual grids that can be applied to observational reference frames in order to identify physical phenomena observable from the standpoint of those frames of reference. There are two main categories of coordinate systems: On the one hand, there are those that are applied “externally” in the sense that the coordinate grids they establish can be thought of as nets consisting of individually identifiable grid intersections that conceptually can be “overlaid” upon physical phenomena given in an observational reference frame, thereby making identifying reference to those phenomena possible. This “overlying” is *extrinsic* and, from a conceptual perspective, *separable* from the objects of reference to which such an external coordinate system is applied. An alternative way of understanding the extrinsic nature of such coordinate systems is to conceive of the objects to which they permit reference as located in an encompassing, usually Euclidean, space. The use of a standard Cartesian coordinate system, for example, allows one to “step back,” so to speak, and consider the objects of reference identified in the coordinate system from the standpoint of the “ambient space” in which they are situated.

Einstein saw that extrinsic coordinates cannot be defined in gravitational fields due to the non-Euclidean nature of such fields and due to the continuous variability of the space-time such fields define. In the interests of as much theoretical simplicity as possible, a system of reference was desirable which is both capable of specifying and characterizing objects and events in a gravitational field, or in a state of acceleration, and which is not applied extrinsically, but is rather “intrinsic”—inseparable from the nature of the field itself. Such a system of reference does not make recourse to the perspective offered by an ambient space, but is rather a system of reference intrinsic to what in this study I’ve called the ‘referential field’ that is in view—here, the referential field is a gravitational field.

Einstein found the basis for such a system of reference in Gaussian coordinates. The insight involved is important and relates directly to fundamental results we have reached in this study. We need a clearer understanding of what is at issue.

## 26.8 Einstein’s mollusks and intrinsic reference

Nearly a century before Einstein succeeded in formulating his general theory, Gauss (1828) published what he called his ‘*Theorema Egregium*’, his “Remarkable Theorem,” which showed that the curvature of a surface can be determined purely by means of measurements that are limited to and made upon the surface itself so that the determination of its curvature does not re-

quire recourse to a space in which the surface is considered to be situated. From the standpoint, for example, of fictitious beings who live on a curved surface, using Gauss's approach it would be possible for them to determine its curvature wholly through simple geometrical measurements made upon the surface itself; they would have no need to leave the surface to observe it from an external—extrinsic—point of view. Gauss's *Theorema Egregium* showed how the curvature of a surface can in this way be determined *intrinsically*; his theorem demonstrated that the resulting "Gaussian curvature" of a surface is an intrinsic and *invariant* property of that surface, independent of situating the surface in an ambient space.<sup>308</sup> The Gaussian curvature, sometimes called 'total curvature',<sup>309</sup> supplies all the information needed to determine the geometry of the surface in question—whether it is, for example, Euclidean, elliptic, or hyperbolic. No matter how the surface is distorted, without stretching, the Gaussian curvature does not change, but is invariant. Einstein recognized the applicability of Gauss's work to gravitational fields.

Einstein seldom made use of metaphors to explain special and general relativity, but he found an interesting and informative metaphor with which to describe one of the essential mathematical tools upon which general relativity is based. The metaphor involves the *mollusk*, an invertebrate member of the phylum that also includes slugs, snails, and octopuses. Mollusks are creatures with soft, pliable, unsegmented bodies. Their bodies are flexible, elastic, and supple, conforming easily to the shape of a surface on which they may attach themselves. Einstein employed the mollusk as a metaphor to communicate how Gaussian coordinates function.<sup>310</sup>

If a set of arbitrary non-intersecting curves is drawn on a surface so that through each point on the surface one and only one curve passes, and another similar set of curves is drawn that overlaps the first set, forming a grid of the intersecting curves of the two sets, then each intersection can be assigned a pair of numbers, the Gaussian coordinates that identify that point. If the axes defined by the curves are  $u$  and  $v$ , then the Gaussian coordinates of a point may be represented a pair of numbers  $(u,v)$ . The grid formed by the two systems of curves conforms to whatever the curvature of the surface may be and makes it possible to specify any individual point on that surface, which may

---

<sup>308</sup> Examples of such Gaussian invariants are the area of a portion of a given surface, another is the length of a curve drawn upon the surface, another is the angle formed by two curves drawn upon the surface that meet at a common point. The modern concept of metric tensor, important in general relativity, evolved from a study of such invariants.

<sup>309</sup> As, for example, by Kreysig (1991, p. 131).

<sup>310</sup> Gaussian coordinates formed the basis for later work in mathematics, where they have come to be called 'generalized' or 'curvilinear coordinates', and sometimes 'synchronous coordinates' as in Stephani and Stewart (1990, p. 20).

be non-Euclidean. Gaussian coordinates, in other words, are *intrinsic* to a given surface. Gaussian coordinates are, so to speak, *embedded in* the surface they describe; such coordinates supply, in the terminology of the present study, information that is “ingredient in them” concerning the curvature they identify, called the ‘metrical properties of the surface’. Einstein used the intrinsic nature of Gaussian coordinates to replace the extrinsic means of reference of the special theory of relativity which, as we’ve seen, employs inertial reference frames. The general principle of relativity could then be stated in the form that “[a]ll Gaussian coordinate systems are essentially equivalent for the formulation of the general laws of nature” (Einstein, 1920/1917a, p. 97, original italics).

As already noted, there are no “rigid bodies” in gravitational fields, and the application of extrinsic Euclidean coordinate systems fails to be of use. A *non-rigid, intrinsic basis of reference* was needed, a means of reference that is compatible with any arbitrary state of motion, and in terms of which time-measurement can be made, subject only to the condition that, at each point on such a non-rigid basis of reference, adjacent time measurements differ infinitesimally little, one to the next.

This non-rigid [basis of] reference, which is not wrongly called “reference mollusk” [*Bezugs-molluske*] is essentially equivalent to any Gaussian four-dimensional coordinate system. What gives the “mollusk” a certain vividness in comparison with the Gaussian coordinate system is the (actually unjustified) formal preservation of the special existence of the spatial coordinates with respect to the time coordinate. Each point of the mollusk is treated as a point in space, every material point at rest relative to it as at rest as long as the mollusk is treated as a reference body. The general principle of relativity demands that all these mollusks with equal rights and equal success can be used as reference bodies in the formulation of the general laws of nature; the laws are considered to be completely independent of the choice of mollusk. (Einstein, 1920/1917b, p. 67, author’s translation)

To make this metaphorically expressed conception meaningful in the reader’s mind requires, to be sure, a certain strength of imagination and intuition. A “reference-mollusk” is, as Einstein admitted, a purely spatial metaphor; it does not provide a way of characterizing the variability of time measurements in a gravitational field. But even with this limitation, Einstein’s mollusk

metaphor emphasizes the reliance of general relativity upon intrinsically based reference. The important and essential point is to recognize that Einstein's objective was to find a way to make wholly intrinsic reference possible within a physical referential field in which spatial lengths, durations, and the simultaneity of events no longer are thought to have stable, univocally determined, framework-independent status.

It has been one of the central goals of this study to shift from the traditional, naive, subject/object-based notion of extrinsic reference, which has separated reference from object of reference and has dominated past philosophical approaches to the theory of reference, to a de-projective concept of intrinsic reference. It is a shift that challenges in many fundamental ways a manner of thinking about and of understanding the nature of reference. Einstein's general theory of relativity accomplished this shift. However, epistemologists and even philosophers of physics have failed to recognize how a parallel shift can—in a manner that is “strongly compelling”<sup>311</sup>—be realized in connection with a general theory of reference.

In making this shift to intrinsic reference, Einstein relied upon additional mathematical tools developed by others in the years following Gauss's publication of his “Remarkable Theorem.” We turn to look at one of these tools that has played a chief and essential role in the development of Einstein's general theory; our interest will focus on the way in which the presupposed mathematics is able to contribute to our understanding of intrinsic reference.

### 26.9 The concepts of tensor and tensor field

It is impossible without mathematics to explain the theory of tensors; the non-mathematician must be content to know that it is the technical method by which we eliminate the conventional element from our measurements and laws, and thus arrive at physical laws which are independent of the observer's point of view. Of this method, Einstein's law of gravitation is the most splendid example.

– Bertrand Russell (1958, p. 90)

The magic of this theory will hardly fail to impose itself on anybody who has truly understood it; it represents a genuine

---

<sup>311</sup> See {11.4} and {17.1}.

triumph of the method of absolute differential calculus, founded by Gauss, Riemann, Ricci, and Levi-Civita.

— Albert Einstein (translated and quoted in Lanczos, 1974, p. 213)

As I think it can be shown, Russell was unduly pessimistic in his claim that non-mathematicians are forever barred from an understanding of the theory of tensors. There are, to be sure, different levels and kinds of understanding. In mathematics, there is the familiar level promoted in our schools whose goal is the cultivation of students' abilities in a given branch of mathematics to use a specialized notation and to perform calculations using it. There is also a meta-level understanding of the most basic concepts a branch of mathematics presupposes; this is the domain of metamathematics and of the philosophy of mathematics. In contrast, in the present study when we discuss the concepts of tensors and tensor fields, the level that will concern us is an understanding of the fundamental referential nature of tensors and tensor fields—that is, their purpose and how they are able to realize it—specifically how and what they make it possible to refer to.

First, by way of brief background:

Einstein recognized, thanks to the work of Riemann, that Gauss's intrinsically defined method of identifying points in the two-dimensional continuum of a surface could be extended to continua of three or more dimensions. Adding to this foundation, today known as intrinsic differential geometry, tensor theory was developed in the late 19th century by Ricci-Curbastro and Levi-Civita. At that time it was called the 'absolute differential calculus' and was later given the modern name 'tensor calculus' as a result of Einstein's work. Together, these contributions, made in successive layers, each depending upon the layer below, formed the eventual mathematical basis for Einstein's general theory of relativity. The result of these contributions, the generalized calculus of tensors, was developed many years before Einstein reached the time when he needed it; to a significant extent the required mathematics was waiting for him.

The word '*tensus*' in Latin means 'taut' or 'tension'; the Latin '*tendere*' means 'to stretch'. In purely general terms, a tensor, as will be explained more fully, is an abstract structure that functionally interrelates a group of components that may be associated with a wide variety of referents. The special advantage that such an abstract, formal structure brings with it is that the functional interrelationships it expresses are translatable, without loss of validity, among any systems of coordinates. This is the chief utility of tensor

theory in the application Einstein made of it—specifically, to provide a method that both is capable of expressing functional interrelationships that may vary from point to point in a space whose geometry may be non-Euclidean, and is able to express such interrelationships in a manner that does not depend on the particular way in which such a space is mapped by a given coordinate system. Let us explore briefly what this entails.

As described earlier in this chapter, if physical objects or events are to be identified through physical measurements, a physical reference frame must be employed, and, in relation to that physical reference frame, an open-ended multiplicity of possible abstract coordinate systems may be employed to designate physical phenomena that may be given in that reference frame. Recourse in this way to a particular system of coordinates provides an abstract identification grid that serves an essentially *auxiliary* purpose, providing what is believed to be a convenient reference system in terms of which a range of phenomena may be identified. The coordinate systems with which we are most familiar—whether Cartesian, polar, spherical, cylindrical, etc.—are, as previously noted, applied extrinsically, allowing one to situate physical phenomena in terms of an ambient space. Doing this worked well enough when inertial systems of reference were involved, but once continua exemplified by gravitational fields are considered, non-extrinsic means of identification, as we've seen, become essential. Since the identification grids of alternative auxiliary coordinate systems comprise merely abstract scaffoldings which we are free to choose, Einstein found that it was possible to express the laws of nature in a coordinate-free form, a form invariant no matter what coordinate system was chosen. This was the objective of the general theory of relativity, and this is where the concept of tensors became crucial.

The subject that will then concern us is how identifying reference to physical phenomena may be possible in a way that is independent of whatever system of coordinates may be used, and in a way that is invariant with respect to possible coordinate systems taken generally.

We have recognized already that the identification and re-identification of the physical phenomena studied by physics require measurement. The concept of tensors presupposes the understanding of two prerequisite kinds of reference to measurements: When measurements are made, for example, of varying temperatures of the surface of a tabletop, the individual temperatures comprise *scalars*—70°F at one point, 73°F at another, etc. A *scalar field* of the tabletop's temperatures has values that may be represented by a single variable that has a specific temperature value at each point on the tabletop. Such measurable individual temperatures are what they are in a manner indifferent to the reference frames of different observers, and irrespective of the

coordinate systems used to identify them. For an observer who is moving and for an observer who is at rest in relation to the tabletop, the temperature measurements of the tabletop's surface remain the same from the two observers' perspectives. When this condition is fulfilled, we may express this fact by saying that the scalars in question are invariant with respect to the transformation of physical reference frames, or of coordinate systems. Scalars with a single component (temperature in this example) that transform in this way are called '*tensors of rank 0*'. But not all scalars are invariant like this: For example, as we know from everyday experience, the frequency of a siren (a scalar quantity) heard by an observer at rest relative to a siren that is stationary will not be the same frequency heard by an observer moving rapidly toward the source of the siren.

The case with *vector* measurements follows in parallel: We recall that vectors have direction and magnitude—for example, the direction and speed of wind at a certain location. A wind map is commonly employed that represents the directions and different speeds of wind at different locations by means of arrows of different lengths; such a map is an example of a *vector field*. Wind speed and direction are two components of such vectors. Vectors may have multiple components.

Vectors that transform invariantly, independently of the observer's reference frame, or of the coordinate system used, are said to be '*tensors of rank 1*'. An example is a vector representing the magnitude and direction of the force at a point in a magnetic field. While such a vector is coordinate-independent, its individual components may not be since they may vary from one observer's reference frame to another, or one coordinate system to another. However, despite such variations of the individual components, the vector quantity itself remains invariant. (It helps to recall that a vector's components in any coordinate system are simply the projections—in the mathematical, not metalogical, sense—of the vector on that coordinate system's axes.)

Bearing the etymology of the word 'tensor' in mind, we find that during its historical development the concept of tensor realized perhaps its most natural and intuitive application in an analysis of the stress of materials. Stress is measured in terms of force per unit area, while stress commonly can be both tensile (normal or perpendicular to a material) and shear (tangential to the material). To express stress that has both magnitude and the two components, tensile and shear stress, requires a *tensor of rank 2*. And there are tensors of successively higher ranks.

We can immediately see that in progressing from scalars, to vectors, to tensors their mathematical representations supply additional information at



each step. The increasing complexity of physical functional interrelationships calls for the representation of an increasing number of “components,” which are the physical, measurable factors the functional interdependencies of which a mathematical physicist seeks to model.

The referential nature of tensors, which concerns us here, can then be described in the following terms: An abstract structure is designed that is capable of representing the functional interrelationships among a number of factors, a tensor’s components, which a physicist judges to be significant. When that abstract structure is able to serve as a formal representation of the relevant functional interrelations in a manner that is independent both of any particular physical reference frame and of the coordinate system used in association with that frame, and when that formal expression is invariant no matter what individual physical reference frame or coordinate system is employed, then the abstract structure that accomplishes this is called a ‘tensor’. A *tensor field*—paralleling the concept of a scalar field, such as a wind map that displays the surface wind velocity on a curved space representing the Earth’s surface—assigns a tensor to each point of the reference space that pertains to the physical phenomenon in view. Examples include stress and strain in a material, which can be expressed by a tensor field, and the tensor that measures the intrinsic curvature of a surface, expressed by the Riemann curvature tensor, which in fact comprises a tensor field.

We now return to Einstein’s reference-mollusks, which served him in communicating in a figurative way the concept of intrinsic Gaussian coordinates. As the theory of general relativity was developed, the place of the mollusk metaphor was taken by a succession of increasingly more sophisticated tensors capable of representing the metric—the intrinsic curvature—of a gravitational continuum: first, the Riemann tensor, which led to the Ricci tensor, and then to the Einstein tensor, details about which cease to be relevant to the focus here of the metatheory of reference.

The purpose of the foregoing explanation is to enable us to see in abstract though clear terms that tensors and tensor fields make it possible to represent functional interrelationships or interdependencies among physical factors judged to be important. They accomplish this by means of formalisms that remain structurally the same, unaffected by variations in the values of a tensor’s components, regardless of changes in observational reference frames, and unaffected by the choice of a coordinate system in terms of which those interrelationships may be expressed.

The resulting degree of comprehensiveness, of theoretical generality, that tensor analysis accomplishes is, as both mathematical and non-mathematical readers should be able to appreciate, conceptually remarkable. Often, tensor

analysts picturesquely refer to tensors as “machines” with “input slots,” “outputs,” “gears,” and “wheels” that “spit out” results when they are fed.<sup>312</sup> Tensors behave—in an admittedly complex way, one not readily accessible to everyday intuition—as operators, taking measurements supplied as inputs, and providing information in return that is coordinate-free and invariant, valid irrespective of the observer’s framework or the particular choice of coordinate system which an observer employs to identify physical phenomena.

Perhaps in this section the veil that has for too long obscured the nature of tensor analysis has in an intelligible manner for non-mathematicians been lifted a little bit. The purpose has not been, as commented early in this chapter, to provide an exposition of special and general relativity, nor in this section to enable the reader to learn the basics of tensor analysis, but rather to describe in abstract terms how Einstein’s two theories establish referential frameworks in terms of which physical phenomena and their functional interdependencies can be represented in a way that identifies physically significant invariances, the laws of nature. In the course of this chapter, we’ve seen how the classical mechanics of pre-relativity physics has been extended by means of the Lorentz transformation, and then how special relativity was extended in the general theory of relativity by means of the intrinsic representation afforded by Gaussian coordinates and then supplemented by tensor analysis. Einstein’s ultimate and successfully realized objective was to characterize physical laws in a manner invariant across physical reference frames, irrespective of the systems of coordinates that may be used to formalize those laws. Throughout the steps we’ve traced, invariance has remained the core concern. Invariance was Einstein’s primary interest in his life as a theoretical physicist, as it is a dominant concern in this study. The concept of invariance and the related concept of covariance stand in need of separate discussion.

### **26.10 Invariance, covariance, and the metalogic of reference**

The laws of nature should be covariant relative to any continuous transformations of coordinates. This postulate (conjointly with the postulate of the greatest possible logical simplicity) limits the laws of nature in a manner incomparably stronger than the principle of special relativity.

– Albert Einstein (1965/1956, p. 174, author’s translation)

---

<sup>312</sup> For example, Misner, Thorne, and Wheeler (1973, §§3.2, 5.1, 9.5 and *passim*), Simmonds (1994/1982, p. x), or Feng (2017, pp. 10, 17).

The terms ‘invariance’ and ‘covariance’ are often used with overlapping meanings, sometimes interchangeably, and sometimes differently by different authors. For clarity, we may distinguish the following meanings: Quantities that are generally considered to remain *constant* irrespective of the individualizing conditions of reference frames may be called ‘invariants’: The speed of light in a vacuum and unaffected by gravitation or acceleration, or the space-time interval between two events, are two such invariants. Related is the sense in which the intrinsic curvature of a surface is considered to be an invariant; Gaussian curvature is such an intrinsic invariant. A second meaning of ‘invariance’ relates to principles or rules that hold true when applied restrictedly or universally to classes of reference frames; this is the major meaning of the concept of invariance we have encountered in describing Einstein’s work. From the standpoint of pre-relativistic classical mechanics, the laws of motion were considered invariant with respect to all inertial reference frames; this kind of invariance is called ‘Galilean invariance’. From the standpoint of the special—the restricted—theory of relativity, the laws of nature are invariant for all inertial systems as prescribed by the Lorentz transformation. From the standpoint of general relativity, the laws of nature are invariant with respect to all reference frames, both inertial and non-inertial.

The term ‘covariance’ can be defined with respect to the *mathematical form* in terms of which rules or principles are expressed. A formulation of a rule or principle is considered to be covariant when its mathematical form remains the same under a specified transformation, or set of transformations. In this sense, Newton’s laws of classical mechanics are covariant under the Galilean transformation; the formulation of physical laws by special relativity is covariant under the Lorentz transformation; the expression of physical laws by general relativity is covariant under any arbitrary (differentiable<sup>313</sup>) coordinate transformation—a comprehensive degree of covariance called ‘general covariance’. The general theory of relativity expresses physical laws in general covariant form: This means, as we’ve seen, that general relativity expresses the laws of nature in a way that is coordinate-independent, irrespective of the choice of coordinate system.

Using the two principal meanings of invariance and covariance together, and from the brief description of tensors in the previous section, we may say that in general relativity tensors represent functional interrelationships that express in general covariant form physical quantities that are invariant, that must appear the same from the standpoint of all observers.

Here, let us take a reflective step back to see more distinctly what is involved from the standpoint of the metalogic of reference. By means of the

---

<sup>313</sup> In general terms, this means continuous at every point in the domain involved.

concept of covariance, by extension we express the *translatability* of an expression whose form is preserved in relation to one or more alternative frames of reference in terms of which that expression can be expressed. Covariance is a mathematical concept, but it need not be restricted to mathematical expressions. The class of mathematical expressions to which the concept of covariance is applied consists most commonly of expressions of functional relationships. Suppose we consider an expression in English; further suppose that the expression comprises an assertion claiming that a functional relationship of a certain kind is the case. The expression when formulated in English has a certain form: Let us say it has the grammatical form of a truth-functional assertion. If the translation of the expression into another natural language retains the same form, we may say that the English expression is covariant in relation to the other language. When translation into more than one other language preserves the form of the original expression, the English expression may be said to be covariant over a correspondingly enlarged class of languages. If, still further, the English expression can be translated into any arbitrarily chosen natural language, without loss of the original expression's form, we may say that the English expression is generally covariant. This simply means that, with respect to natural languages considered generally, the English expression can be translated without the loss of its original grammatical form as a truth-functional assertion.

We've seen that invariance in relativity physics relates to the constancy of the empirical truth of the fundamental propositions of special and general relativity, a constancy that is preserved, respectively, by the Lorentz transformation among inertial systems of reference, and by arbitrary coordinate transformation (i.e., general covariance) with respect to both inertial and non-inertial systems of reference.

The level of maximum theoretical generality on which the metalogic of reference is established requires that the expression of its results has such a degree of invariance: Those results must hold irrespective of the frame of reference to which they are applied. In fact, the invariance required is somewhat more than this since the results reached must not only be translatable without loss of truth to any arbitrarily chosen frame of reference, but those results must be equally true when applied to the meta-framework presupposed by the metalogic of reference itself. The invariant results of the metalogic of reference are, in the language of relativity theory, generally covariant, as well as being self-applicable. Where the purpose of Einstein's special and general theories of relativity is to formulate laws of nature that hold invariantly, irrespective of the conditions of individual observational frames of reference, the purpose of the metalogic of reference is to recognize and formulate universal

principles of invariance that govern the referential capacities of all frames of reference, irrespective of the ends to which they are employed or the sets of the objects to which they provide the basis of reference.

### **26.11 The convergence of relativity physics and the metalogic of reference**

Einstein's success in developing the special and general theories of relativity can be understood in terms of the acuity of his analysis of the conditions of macrophysical reference; later we shall similarly discuss quantum theory's analysis of the conditions of microphysical reference. From the standpoint of the present study, the success of the special theory can be understood to have resulted from Einstein's realization—which, as we've pointed out, he recognized as a supposition—that the transmission and reception of light signals of constant velocity in a vacuum are required for physical identifying reference to the simultaneity of events, to measurements of durations, and to lengths. The success of the general theory rested on his recognition that the mathematical expression of identifying reference to physical phenomena that are subject to acceleration or gravitational fields requires the intrinsic referential nature of tensors. Both the special and the general theory of relativity may be understood as based on an implicit, sometimes perhaps only intuited, epistemological understanding of the conditions of physical reference from the standpoint of inertial and non-inertial observational reference frames.

A number of important results of the special and general theories converge with results we have reached in this study by means of an altogether different, non-empirical, non-physical route. The following table compares a group of these frequently parallel results and the corresponding methodological principles that define each approach:

RESULTS REACHED AND METHODS USED BY RELATIVITY PHYSICS	RESULTS REACHED AND METHODS USED BY THE METALOGIC OF REFERENCE
<b>PART I</b>	
<i>Rejection of the following pre-relativity claims and notions:</i>	<i>Rejection of the following traditional philosophical and commonsense claims and notions:</i>
(a) “The time-interval (time) between two events is independent of the condition of motion of the body of reference” (Einstein, 1920/1917a, p. 30).	Projections of absolute time; also projections of the past, of time-flow, and of the future ({{22}}).
(b) “The space-interval (distance) between two points of a rigid body is independent of the condition of motion of the body of reference” (Einstein, 1920/1917a, p. 30).	Projections of absolute space; also projections of space-time ({{22}}).
(c) “The simultaneity of two events determined from the standpoint of an inertial system implies the simultaneity of the events in relation to every inertial system” (Einstein, 1965/1956, p. 170, author’s translation).	Time specifications have no possible meaning independently of the frame(s) of reference in terms of which they are made ({{22}}).
(d) There is no such thing as absolute simultaneity: “it is necessary to abandon the absolute character of simultaneity” (Einstein, 1965/1956, p. 169, author’s translation).	

<p>(e) “[T]here is no such thing as an independently existing trajectory ... , but only a trajectory relative to a particular body of reference” (Einstein, 1920/1917a, p. 10)</p>	<p>Projections of things-in-themselves ({21}).</p>
<p>(f) There is no such thing as “motion in itself, to which we can attach no meaning” (Einstein, 1920/1917b, p. 37, author’s translation)</p>	
<p>(g) There is no stable, unchangeable inertial basis of reference in itself: “the fictitious rigid body of reference is of no avail in the general theory of relativity” (Einstein, 1920/1917a, p. 99)</p>	
<p>(h) There is no “empty space” in itself. “If one supposes that the field of gravitation is eliminated, then there remains “absolutely <i>nothing</i>, not even a ‘topological space’.... [An] empty space, that is a space without a [gravitational] field, does not exist” (Einstein, 1920/1917b, pp. 177-178, author’s translation).</p>	
<p>(i) There is no preferred, privileged observational frame of reference.</p>	<p>Projections of the self as center of experience ({24.6}).</p>

<b>PART II</b>	
<i>Methodological principles that define each approach:</i>	
(j) The representation and formulation of physical laws must accept the framework relativity of physical phenomena.	The representation and formulation of the principles that govern the identifiability of objects must accept the objects' framework relativity.
(k) The physical identity of phenomena—i.e., their measurable properties—is essentially a function of the reference frame employed to identify them.	The ontological identity (or status) of objects of reference—i.e., their identity as entities to which reference is possible—is essentially a function of possible reference frames in terms of which they can be identified.
(l) Laws of nature are formulated in a form that is covariant—i.e., independent of any particular coordinate system—and empirically invariant—i.e., true irrespective of the conditions of particular observational reference frames.	The principles governing possible reference are formulated in a covariant manner—i.e., translatable without loss of validity among arbitrarily chosen frameworks of reference—and these principles are invariant—i.e., they cannot <i>not</i> be accepted without undermining their own possible meaning
(m) Relativity theory specifies physical conditions of measurement essential to the identification of physical phenomena.	The metalogic of reference specifies abstract preconditions of identifying reference with respect to any objects of reference.



<p>(n) Relativity theory shifts from the use of extrinsic pre-relativity frameworks of reference to intrinsic mathematical means of identification drawn from Gauss and Riemann.</p>	<p>The metalogic of reference shifts from the use of extrinsic, traditional relational notions of reference to an intrinsic understanding of identification.</p>
<p>(o) Relativity theory applies intrinsic means of mathematical representation and formulation of functional interrelationships of physical phenomena, using tensor analysis.</p>	<p>The metalogic of reference applies intrinsic means of abstract representation and formulation of the concept of embedment of preconditions of reference in the identity of objects, using the concept of referential fields.</p>
<p>(p) The approach is proposed as “a valuable heuristic aid in the search for general laws of nature” (Einstein, 1920/1917a, p. 43).</p>	<p>The approach is developed in terms of heuristic tools of analysis ({14.4}).</p>
<p><b>PART III</b></p>	
<p><i>Major differences between the two approaches:</i></p>	
<p>(q) Relativity theory seeks to formulate a comprehensive physical theory that is as conceptually and mathematically simple as possible, from which empirically confirmable laws of nature can be logically derived.</p>	<p>The metalogic of reference seeks to formulate a comprehensive metatheory of reference that cannot <i>not</i> be accepted without undermining the possibility of meaning—i.e., an undeniable general theory of identification and meaning.</p>

**Table 26.1 Comparisons of relativity physics and the metalogic of reference**

A few comments may be made about entries in the preceding table: Einstein's results (a) – (f) and (i), which rejected fundamental tenets of Newtonian physics, were among the direct consequences of the special theory of relativity. All of these tenets, which were believed for centuries, had to be relinquished once the constancy of the speed of light in a gravity-free vacuum was accepted as a physical condition of the measurement of—i.e., identifying reference to—times and distances. The notions of absolute space; absolute time, universality of the uniform measurement of time, and the simultaneity of events; the putatively meaningful framework-independent “in-themselves” of such physical phenomena as motion and the trajectories of objects; and the notion of a fixed, stable, unchangeable, privileged inertial frame of reference—all had to be given up when Einstein showed that they lead to unacceptable consequences. Among the results of the general theory of relativity, only (g) and (h) have been selected for inclusion in Part I of the table, while, in Part II, (j), (k), and (m) played fundamental roles in both special and general relativity; (l), (n), and (o) identify breakthrough advances made by general relativity; and (p) classified relativity theory as a “heuristic aid,” a guide to problem-solving, not a petrified set of doctrines.

As the table summarizes, a number of the central results of relativity physics and the methodological principles followed by Einstein in developing the special and general theories have clear parallels both with results we have reached in this study and with methods we have used in the process. Since both approaches are firmly grounded in a purely theoretically based recognition of framework relativity, the parallels noted in the table should come as no surprise. Indeed, *a number of the main results of this study establish and confirm on a metalogical level corresponding principal results of special and general relativity.*

And yet, despite the parallels, the two approaches could not be more different: the one belongs to theoretical physics, the other to a non-physical meta-framework designed for pure conceptual analysis. Where Einstein's contributions are evaluated by physicists in terms of the physical confirmability or falsifiability of the consequences they predict, any relevant assessment of the metalogic of reference must be situated on a wholly abstract, epistemological level. Most importantly, there are these differences: The metalogic of reference studies the limits, the horizons, of possibility and meaning; the metalogical horizons it serves to identify constitute limits of what is possible and of what is thereby *possibly meaningful*; the “modal level” of its analyses is poles apart from the empirical, factual world of physics; and, most importantly, its conclusions follow necessarily because they cannot be denied without undermining the very possibility of the meaning of such denial. The latter

property of the metalogic of reference is, of course, far removed from the defining properties of results in theoretical physics in which—to mention but an obvious example—there is nothing either inherently logically self-contradictory or metalogically self-undermining in denying that the speed of light in a gravity-free vacuum is *not* a constant. The latter is an issue of physics, not of the metalogic of reference.

As we've seen in the course of this chapter, Einstein made a number of interrelated claims in which he asserted that certain central and basic pre-relativity notions were *meaningless*—that time specifications have no meaning unless associated with suitable frames of reference, that the length of physical objects has no meaning independently of the frames of reference employed in their measurement, that simultaneity has no meaning unless based on an empirical decision method that can determine whether a pair of events is simultaneous, that “motion in itself” has no meaning, etc. If we ask what justification Einstein had for such indictments of meaninglessness, we are forced to read between the lines and make recourse to imagination: He may have meant that these notions are inherently framework-relative because this fact should be clear and obvious in light of the convincingness of the special and general theories of relativity, and that, hence, these notions should be dismissed as having, we might say, “no physical meaning”—that is, within the context of relativity theory. Or perhaps he meant that his special and general theories of relativity constitute major advances over pre-relativity physics, so much so that his criticism that the absolutist classical notions have “no meaning” may actually have been simply an expression of justified, impatient dismissal. Still another possibility that might occur to one is that he adhered to an operationalist criterion of meaning, which, if applied to the above notions, could lead to the conclusion that they are meaningless. However, we saw earlier in this chapter that Einstein rejected operationalism (“it is not necessary to demand that all of [physics'] assertions can be independently interpreted and ‘tested’ ‘operationally’; *de facto* this has never yet been achieved by any theory and can not at all be achieved”).

In short, it is not clear what, if any, justification Einstein may have had in mind to support his contentions regarding the “meaninglessness” of such absolutist notions as time specification, simultaneity, length, and motion. However, here, as in the case of a wide range of other projective notions and claims, the metalogic of reference supplies the missing justification, and with a scope of application that goes far beyond theoretical physics, applying universally to any framework of reference that permits identifying reference to a range of objects.

Despite these major and fundamental differences both in subject-matter,

basis of justification, and range of application, the special and general theories of relativity provide, to my knowledge, the best examples to date of any discipline's contributions that have come about as a result of careful, disciplined, steadfast respect for and adherence to framework relativity. The consequences to which special and general relativity have led have been revolutionary and have forced the abandonment of a number of deeply rooted classical presumptions and notions that were not easy for many physicists to relinquish. In a multitude of ways, these consequences have seemed counterintuitive and hard to incorporate in the thinking and outlook of more than a limited group of specialists. And yet, with the passage of time—now, more than a century—relativity physics has become a familiar topic even to children. Its strangeness has ebbed away; the revolution in thinking it inspired has now largely been absorbed and accepted.

In a corresponding way, the consequences of the metalogic of reference evidently pose a challenge to traditional philosophy. Those consequences are often counterintuitive, rejecting traditional philosophical assumptions and customarily agreed upon notions that have, for millennia, been employed unquestioningly as meaningful. Any approach whose results have these properties is revolutionary and is therefore unlikely to be easily accepted.<sup>314</sup> When a new paradigm, a new research program, is proposed, it seems—given the present general human constitution and the characteristic contentiousness of much philosophy as described in the first two chapters of this study—that it cannot be otherwise.

The convergence of relativity physics and the metalogic of reference described in this section is, as I see this, far from accidental, but reflects a shared concern to abide by and to apply principles of framework relativity in a manner that is capable of identifying invariants in the two domains involved, theoretical physics and the theory of reference. Despite this convergence, Einstein's commitment to the principles of framework relativity weakened very considerably over the years, and then, and there is no other honest word for it, that commitment *dissolved*—paradoxically and ironically—when he was later confronted by the results of quantum theory, as we shall see in the next chapter.

---

<sup>314</sup> The author's experience, over a period of more than 50 years, has confirmed that this is the case. A brief autobiographical account is given in Bartlett (2017a).

## Quantum Theory as Seen through the Lens of the Metalogic of Reference

### 27.1 Introductory comments

As in the preceding chapter's discussion of the physics of relativity, it will not be my purpose in this chapter to provide a general exposition of quantum theory, to discuss individual experimental results, or to analyze the variety of alleged paradoxes with which some physicists, many philosophers of science, and many popularizers of science have become concerned.<sup>315</sup> Nonetheless, this chapter will not assume that the reader is familiar with the basic, now strongly established, and widely accepted conclusions of quantum theory, which in the course of the chapter will be identified and briefly reviewed. But a deeper and more detailed familiarity with quantum theory's results is undeniably helpful in order more fully to appreciate the very significant step that quantum physics has made in understanding physical reality.

If you are a reader especially interested in quantum theory, but less so in the special and general theories of relativity, and you have skipped the preceding chapter and come directly to this one, you will find it useful to read {26.1}. Most relevant to the subsequent analysis of quantum theory are the following considerations which were discussed there: the transformability of the way in which an object of reference is identified in one, or a group, of reference frames so that through such a transformation its identification conforms to the requirements of one or more alternative systems of reference; the distinction between compatible and incompatible reference frames; the concept of identifying reference as it is represented in its more specialized physical variety; the physical basis of frames of reference that provide the means to determine, in relation to them, the spatial and temporal orientation of the physical phenomena studied; and the role of alternative coordinate systems in terms of which it is possible to represent or designate a physical phenomenon

---

<sup>315</sup> There are numerous good introductions to quantum theory; many of the works by Heisenberg and by Bohr are accessible to the non-physicist and are very clearly written; some of these are included in the References at the end of this book.

given in a specific physical reference frame. It will also be of value for readers who have skipped directly to the discussion here to review {26.2}.

Our main interest in connection with quantum theory, as was the case with the special and general theories of relativity, will relate to the ways in which the theory exemplifies deliberate, self-conscious, and critical analysis concerning framework relativity. Our focus will be the route taken by quantum theory to provide a specialized, theoretically based system of identification of microphysical phenomena. In the process, we shall seek to make clear certain of the principal referential preconditions that are presupposed, which serve in a limitative manner to define quantum reality.

### 27.2 Measurement-based perturbation

The people of Tlön are taught that the act of counting modifies the amount counted, turning indefinites into definites.

– Jorge Luis Borges (1998/1941, p. 76)

Does measurement “change” what is measured? Certainly in some instances it is very evident that it does. The height measuring device in doctors’ offices can sometimes clumsily be pushed down very firmly on the top of the patient’s head, perceptibly affecting the individual’s height before a measurement is taken. Or consider a soap bubble whose diameter we wish to measure. If our only measurement device is a special pair of large calipers which must be in physical contact with what is measured and which must apply some force to obtain a measurement, it is not unlikely that the soap bubble will pop as it is measured. Such measurements may change not only what is measured, but may destroy the object whose properties are to be measured. Quantum theory appears to bring one face-to-face with this phenomenon of perturbation that can come about when measurements are made. I say “appears” because the question whether measurement “changes” what is measured becomes a good deal more complex and in need of analysis when dealing with the very small scale on which quantum events take place.

When measuring, for example, the size of very minute objects, we require a measuring device with a scale that is fine enough to register, for example, the lengths of the objects. If a rigid ruler can appropriately be used, its calibrations must be sufficiently precise—the size of the objects to be measured cannot be smaller than the distance between the ruler’s calibrations. This observation applies on the atomic scale, as Heisenberg described in his famous thought example of a “microscope” capable of measuring the properties of an

electron. To gain information about its position, ordinary light cannot be relied on

... since the inaccuracy of the measurement of the [electron's] position can never be smaller than the wave length of the light. But a microscope using  $\gamma$ -rays with a wave length smaller than the size of the atom would do.... The position of the electron will be known with an accuracy given by the wave length of the  $\gamma$ -ray. The electron may have been practically at rest before the observation. But in the act of observation at least one light quantum of the  $\gamma$ -ray must have ... first ... been deflected by the electron. Therefore, the electron has been pushed by the light quantum, it has changed its momentum and its velocity.... (Heisenberg, 1958, p. 47)

Two important claims are made in this passage, and they are claims that have been re-asserted many times by other quantum theorists since Heisenberg wrote the above lines. They are explicitly epistemological claims, and their possible meaning requires that specific referential preconditions must be satisfied in order for relevant identifying reference to be possible. Those claims are: (1) the “act of observation” “must have first done something to the object to be measured,” and (2) this “something” produced a “change” in the object to be measured. As F. S. C. Northrup expressed this notion concisely, “The very act of observing alters the object being observed when its quantum numbers are small” (Heisenberg, 1958, p. 24). These two claims have been fused together into the notion that measurement involves “perturbing interaction” between the observer, his measuring apparatus, and the quantum phenomena that are measured.

Among quantum physicists, Niels Bohr was one of the most restrained in his willingness to make claims that exceed the referential capabilities of quantum theory. He sometimes expressed the need for caution in making statements about such “perturbing interactions” when quantum measurements are made. For example, we find the following warning in one of his books:

[O]ne sometimes speaks of “disturbance of phenomena by observation” or “creation of physical attributes to atomic objects by measurements.” Such phrases, however, are apt to cause confusion, since words like phenomena and observation, just as attributes and measurements, are here used in a way incompatible with common language and practical defi-

dition. On the lines of objective description, it is indeed more appropriate to use the word phenomenon to refer only to observations obtained under circumstances whose description includes an account of the whole experimental arrangement. (Bohr, 1958, p. 73)

The problem that is posed, however, does not involve merely a matter of incompatibility with common language and of practical definition, but, as we shall see shortly, a matter of referential impossibility.

Bohr recognized that, from the quantum-theoretical standpoint, the putatively meaningful notion of “perturbing interaction” has “revealed an unsuspected limitation” (Bohr, 1958, p. 74) in the quantum physicist’s understanding of the microphysical world. We shall call this a ‘*limitative result*’ of quantum theory, a result which, as long as quantum theory in its present form is accepted, *in principle* rules out the referential capacity to obtain information about the alleged “independent properties” of the small-scale objects under study. This is not a matter that can be remedied through the invention of more sophisticated and sensitive means of detection; it is not a matter of mere practical unfeasibility at this time; it is a matter of *impossibility in principle*. As Bohr expressed this,

[N]o result of an experiment concerning a phenomenon which, in principle, lies outside the range of classical physics can be interpreted as giving information about independent properties of the objects, but is inherently connected with a definite situation in the description of which the measuring instruments interacting with the objects also enter essentially. (Bohr, 1958, p. 26)

This in-principle-impossibility relates not only to the possible meaningfulness of ascribing independent status to properties of observed quantum phenomena, but it also relates to the possible meaningfulness of causal claims concerning the “influence” of measuring instruments upon the small-scale objects measured. More than these, the impossibility at issue concerns the possible meaningfulness of the compound notion of “perturbation interaction” with which we began, along with the possible meaningfulness of clearly distinguishing quantum objects from the observer and the observer’s measuring instruments in terms of which they are studied and understood.

Before we proceed to consider these mutually entangled problems, let us return to the previous simple, real-world examples of a person’s height



measurement and the measurement of the diameter of a soap bubble. These two examples serve “to keep our feet on the ground” by making clear in concrete instances what referential preconditions are, in fact, satisfied when a claim is made that these measurements “perturb” what is measured. In the first case, the patient whose height is being measured is able to perceive, through proprioception and perhaps also vision if there is a mirror handy, the compression of his or her height when the height lever is pressed down unduly hard on the patient’s head. In the second case, when the soap bubble’s diameter is measured with our imagined calipers, it is possible to see the bubble pop when the measuring calipers touch and exert some force on its film. In other words, in both cases it is possible, in principle as well as in fact, to observe in some fashion a state *before and after measurement*; the observed difference between the two states makes it possible to refer to a *change* that has occurred between the earlier state and the later state. In both cases, the earlier and the later state are continuous in time; we are able, should we wish, to discriminate very small time intervals, as, for example, between the time the calipers make contact with the soap film and the popping of the bubble shortly thereafter. Relevant referential preconditions are satisfied that are necessary in order meaningfully to assert that a change has followed the measurement. (We carefully avoid introducing conventional causal notions in these “measurement perturbation” claims, in accordance with the results of {23}.) In the subsequent sections, we shall see how, in principle, quantum theory rules out that such referential preconditions can be satisfied.

### 27.3 Eliminating “interpretation” from the Copenhagen interpretation

I express the opinion that the Copenhagen interpretation is correct and indispensable. But I have to add that the interpretation has, in my view, never been fully clarified. It needs an interpretation itself, and only this interpretation will be its defense.

– C. F. von Weizsäcker (1980/1971, p. 181)

It may be helpful to state in advance the conclusion we shall reach: Further “interpretation” of the “Copenhagen interpretation” is neither necessary, nor is any “interpretation” whatsoever called for. Instead, the *set of results* associated with the “Copenhagen interpretation” of quantum theory is *unavoidable* in the context of present quantum theory. The understanding of quantum

phenomena which has become known as the Copenhagen interpretation is not itself an actual “interpretation” nor does it require one—if by the term ‘interpretation’ one means, as I shall mean here, the notion that one particular explanation among others is judged to be more *satisfying*, but is not proved to be such that it cannot rationally be rejected. An “interpretation” in this sense is a way of construing a subject-matter that meets certain criteria of taste or satisfaction, but can nonetheless be rejected without incoherence. In my judgment and in agreement with Heisenberg, the variant “interpretations” that have in the past been proposed in connection with what I will henceforth call ‘*the Copenhagen results*’ are precisely of this kind: These “alternative interpretations” do not compel assent because they do not provide justification which demonstrates that any one of them cannot be rejected. In contrast, the Copenhagen results are themselves *physically as well as theoretically inescapable*, a general and universal claim made, though not in these words, by both Heisenberg and Bohr. Equally, and from a theoretically fundamental perspective, more importantly, they are results which, as I shall try to make clear, are *conceptually necessary*, compelling assent from the metatheoretical standpoint of the metalogic of reference. If we can provide justification for this position, we will have responded to von Weizsäcker’s request for the missing “defense” of the “Copenhagen interpretation.”

In 1930, Heisenberg called the Copenhagen results the ‘*Kopenhagener Geist* [spirit or mind] *der Quantentheorie*’<sup>316</sup>—despite the fact that the German words ‘*Interpretation*’, ‘*Deutung*’, and ‘*Auslegung*’ were available had he wished to emphasize the tentative, inconclusive, personal-opinion connotation of the word ‘interpretation’. ‘Interpretation’ in the conventional sense in which I use it here should, one would think, have no place in any rigorous science.

Heisenberg himself seems to have been responsible for coining the phrase ‘the Copenhagen interpretation’, a choice he later came very much to regret—much like Einstein’s later dissatisfaction with his decision to use the term ‘relativity’ to name his special and general theories. In 1955, Heisenberg published a paper, “The Development of the Interpretation of Quantum Theory” (Heisenberg, 1955). In this paper, Heisenberg began to use the ill-chosen phrase. In the same year, he delivered two Gifford Lectures at Saint Andrews University; their titles also incorporated the newly minted phrase: “The Copenhagen Interpretation of Quantum Theory” and “Criticism and Counterproposals to the Copenhagen Interpretation of Quantum Theory”; both papers are included in Heisenberg (1958). Before copies of the book had even been distributed, Heisenberg expressed chagrin that he had used the phrase:

---

<sup>316</sup> Heisenberg (1930, on the last, unnumbered, page of his Preface).

I avow that the term ‘Copenhagen interpretation’ is not happy since it could suggest that there are other interpretations.... We agree, of course, that the other interpretations are *non-sense*, and I believe that this is clear in my book, and in previous papers. Anyway, I cannot now, unfortunately, change the book since the printing began enough time ago.<sup>317</sup>

The word ‘interpretation’ does certainly suggest the possibility of perhaps equally legitimate, alternative “interpretations”—which in turn implies that the Copenhagen results are not demonstrated results at all, but have the status merely of statements that it may be possible to construe in conflicting ways, or evade or reject.

In itself, the unwise naming of a theory can sometimes encourage and open the way for a seemingly endless proliferation of seemingly competitive, alternative rivals. This has indeed happened in connection with quantum theory, as Mermin (2012, p. 8) remarked: “[T]oday, nearly 90 years after its formulation, disagreement about the meaning of the theory is stronger than ever. New interpretations appear every year. None ever disappear.” The parade of “interpretations” continues to lengthen. There are now more than a dozen often conflicting interpretations of quantum theory vying for attention, including the stochastic, von Neumann-Wigner, de Broglie-Bohm, ensemble, many worlds, consistent histories, relational, transactional, objective collapse, many minds, quantum Bayesianism (AKA Qbism), and the list is ever made longer as many of the proposed interpretations have led to their own variations. Again, one cannot but wonder that the results of a rigorous science should stand in need of so much “interpreting.” Since many of these interpretations are implicitly or explicitly motivated by differences of opinion in connection with the central epistemological issues with which we are concerned, we shall leave this clamorous multitude to one side, and return to the conceptual foundation.

Beginning in 1925, physicists Niels Bohr, Werner Heisenberg, Max Born, and others defended the Copenhagen results. These results became the central topic, discussed and argued, at the now famous conference held in Brussels in 1927 by the Institut International de Physique Solvay. The conference was chaired by Hendrik Lorentz, and was attended by many leading physicists, including Bohr, Heisenberg, Born, Schrödinger, Pauli, Dirac, de Broglie, Planck, Einstein, and others. Among the principal assertions made by the Copenhagen results is that any information obtainable from quantum phe-

---

<sup>317</sup> Letter to Léon Rosenfeld dated April 16, 1958, in the Rosenfeld Papers, Niels Bohr Archive, Copenhagen. Quoted in Freire (2005, p. 28, italics added).

nomena is necessarily and essentially related to the instrumentation used by observers to make measurements of those phenomena. It is inextricably woven into the Copenhagen results that it is *without meaning* to ascribe properties to an isolated quantum phenomenon since such a phenomenon cannot be separated from its functional relation to the dynamic system comprised of observer–measurement apparatus–quantum-object. In keeping with the methodological approach of the present study, we’ll refer to this result as *the framework relativity of quantum phenomena*.

The Copenhagen results furthermore showed that there is no *meaning* that can be associated with the notion that a quantum phenomenon—an electron, for example—“in reality” possesses a well-defined position or momentum independently of the precise measurement of the other. If one of these two so-called ‘canonically conjugate’ properties of the electron has been measured—for example, its momentum—then it is physically without meaning at the same time to ascribe to the electron a definite position. Heisenberg called this well-known result ‘*the principle of indeterminacy*’ or ‘*uncertainty*’. Other conjugate properties—for non-mathematicians, these properties are more informatively sometimes called ‘complementary’ or ‘incompatible observables’—include time/energy, angular momentum/angular position, number/phase, etc. The determinate measurement of one member of a conjugate pair brings with it a simultaneous corresponding indeterminacy of the other. The wave/particle duality of light is one of the best-known examples of mutually exclusive properties that cannot both, at the same time and using the same experimental setup, be observed.

We will return to these two fundamental results of quantum theory shortly.

#### 27.4 The Einstein, Podolsky, Rosen (EPR) position

[T]he finite interaction between object and measuring agencies conditioned by the very existence of the quantum of action<sup>318</sup> entails—because of the impossibility of controlling the reaction of the object on the measuring instruments if these are to serve their purpose—the necessity of a final renunciation of the classical ideal of causality and a radical revision of our attitude towards the problem of physical reality. In fact, ... a criterion of reality like that proposed by [Einstein, Podolsky, and Rosen] contains—however cautious its for-

---

<sup>318</sup> I.e., Planck’s constant.

mulation may appear—an essential *ambiguity* when it is applied to the actual problems with which we are here concerned.

– Niels Bohr (1935, p. 697, italics added)

As will be made clear, Bohr’s use of the word ‘ambiguity’ was not only an expression of his diplomacy toward those with whom he disagreed, it was misplaced, for there is no “ambiguity” involved, but rather a self-undermining expression of projective meaninglessness.

The now-classical paper by Albert Einstein, Boris Podolsky, and Nathan Rosen (1935), quoted and briefly discussed earlier in {19.3}, opposed the Copenhagen results by stating, but failing to justify, the authors’ shared *preference in believing* that quantum phenomena possess an autonomous reality, independently both of the context established by individual observations and of the conceptual structure of quantum theory itself. The position they sought to defend expressed a stubborn and indeed unyielding desire to retain the conventional realism of classical physics despite the repetitively confirmed and firmly established Copenhagen results—specifically the “impregnability”<sup>319</sup> of the framework relativity of quantum phenomena and the principle of indeterminacy. Heisenberg recognized that, at its core, the EPR position expressed the *preferred belief* of the three authors—that is, their *wish to believe* in “the objective reality, which is independent of any theory” (Einstein, Podolsky, Rosen, 1935, p. 777). Heisenberg saw that:

They would *prefer* to come back to the idea of an objective real world whose smallest parts exist objectively in the same sense as stones or trees exist, independently of whether or not we observe them.... This, however, is impossible or at least not entirely possible because of the nature of atomic phenomena.... It cannot be our task to formulate *wishes* as to how the atomic phenomena should be; our task can only be to understand them. (Heisenberg, 1958, p. 129, italics added)

Einstein repeatedly expressed this wish in his later publications, the desire for “the complete description of any (individual) real situation (as it supposedly exists irrespective of any act of observation or substantiation)” (Einstein,

---

<sup>319</sup> I re-use the term employed by Heelan (2017, p. 61) in recounting Heisenberg’s recollection “that by the end of 1927, it began to be said everywhere that those people in Copenhagen seemed by all accounts to have an impregnable position....”

1959/1949a, p. 667). We have encountered this wish numerous times and in numerous contexts in earlier chapters, a wish that takes the form of the philosophical and psychological compulsion to trespass beyond horizon boundaries ({20}). It is an inherently projective wish that expresses the compulsion discussed earlier ({14}) in connection with Kant's recognition of a "feeling of being forced [*notgedrungen*] ... to seek ... contentment beyond all the concepts which [the believer] can vindicate by experience" (Kant & Beck, 1950/1783, §57).

The EPR position, as we can now see with the clarity made possible by the passage of many decades, reduces to what is no more than an attempt to persuade others to share in Einstein's, Podolsky's, and Rosen's preferred belief in naive realism as applied to the quantum domain. Einstein's steadfast unwillingness during the remaining years of his life to relinquish and renounce a realism rendered outdated by progress in quantum theory brought his work in theoretical physics to a frustrating dead-end.

It should not be necessary to reiterate the steps of previous chapters ({21–23}) that led to the conclusion that realism is metalogically self-undermining, that realist longings and the claims they give rise to are metalogically self-undermining and therefore meaningless. At this point, we are in a position to be able to place, without misgiving or remorse, a long-overdue tombstone on the grave of the projective, putatively meaningful wish for framework-independent reality, whether on the level of macroscopic *or* atomic phenomena.

Quantum theory is therefore not in any meaningful sense "incomplete," as the EPR position claimed. The notion of the "incompleteness" of quantum theory, presumed in the realist view favored by Einstein, Podolsky, and Rosen to have meaning, undermines the very *possibility* of its meaning. Information obtained by all the possible means afforded by the referential framework of the theoretical system of quantum mechanics and by the experimental confirmation of quantum theory *is complete information*: "Information beyond the theory," which might fill the presumably meaningful "gaps" in our knowledge, beyond the experimental results that can substantiate that theory, is, from the referential standpoint which these establish, an impossibility in principle. Quantum theory as understood in terms of the Copenhagen results is complete, for it includes all that is possible from that standpoint.

In terms of that conceptual framework as it stands today, it is impossible, when describing a conjugate pair of properties of a quantum mechanical system, simultaneously to ascribe definite values to both members of the pair. Einstein, Podolsky, and Rosen believed that quantum theory must be "incomplete" because the impossibility of ascribing determinate simultaneous values to both members of such a pair must, they thought, reflect *ignorance*—the

theory must be “incomplete” because they felt the lack of information they desired, and when such information is lacking, this must mean, they thought, that our knowledge is “incomplete.”

But these are implications in appearance only. They express the demand that a theory of quantum reality should be developed which would *not* presuppose the ways that in principle are possible to refer to it. This demand, which is camouflaged in the EPR assertion of “incompleteness,” cannot be met, again in principle, within the existing theoretical structure of quantum theory. Bohr was surely unaware of this study’s explicitly developed meta-logical basis for the claim that the “completeness” which Einstein, Podolsky, and Rosen wished for is projectively meaningless; instead, Bohr saw their claim as fundamentally unreasonable or irrational: To advance beyond the EPR wish for “completeness,” Bohr developed the concept of “complementarity” in terms of which quantum theory’s descriptions of phenomena “fulfill, within its scope, all *rational* demands of completeness” (Bohr, 1935, p. 696, italics added). We’ll look at this concept in more detail later.

At the beginning of this section, I quoted a passage from Bohr in which he claimed that the EPR position went wrong due to an “essential ambiguity” in their authors’ thinking. Several pages later, Bohr returned to his claim that “ambiguity” was involved:

[W]e now see that the wording of the above-mentioned criterion of physical reality proposed by Einstein, Podolsky and Rosen contains an *ambiguity* as regards the meaning of the expression “without in any way disturbing a system.” ... [T]here is essentially the question of *an influence on the very conditions which define the possible types of predictions regarding the future behavior of the system*. Since these conditions constitute an inherent element of the description of any phenomenon to which the term “physical reality” can be properly attached, we see that the argumentation of the mentioned authors does not justify their conclusion that quantum-mechanical description is essentially incomplete. (Bohr, 1935, p. 700, first italics added)

The fault in Einstein, Podolsky, and Rosen’s reasoning, however, was not due to the presence of genuine “ambiguity,” but rather it came about through self-undermining thinking on a level that concerns the very referential possibility of determining, in principle, the properties of quantum phenomena.

### 27.5 Hidden variable proposals<sup>320</sup>

[A]ll of the opponents of the Copenhagen interpretation do agree on one point. It would, in their view, be desirable to return to the reality concept of classical physics or, more generally expressed, to the ontology of materialism: that is, to the idea of an objective real world, whose smallest parts exist objectively in the same way as stones and trees, independently of whether or not we observe them.

– Werner Heisenberg (1955, p. 17)

One group of counterproposals [to the Copenhagen results] works with the idea of “hidden parameters.” Since the quantum-theoretical laws determine in general the results of an experiment only statistically, one would from the classical standpoint be inclined to think that there exist some hidden parameters which escape observation in any ordinary experiment but which determine the outcome to the experiment in the normal causal way.

– Werner Heisenberg (1958, p. 130)

There has, as one might expect, been strong opposition among realist philosophers to the Copenhagen results. Among the majority of quantum physicists, however, the Copenhagen results, since the time of Bohr and Heisenberg’s work, have been accepted as a well-established basis for progress in theoretical and experimental research in quantum theory. Nonetheless, contrary to this general consensus among quantum physicists, a bias in favor of realism and physical determinism has been expressed by a minority, proponents of so-called ‘hidden variables’. Numerous philosophers and a few physicists have claimed, despite the uncertainty relations, that a microparticle “in fact”—i.e., “in reality”—possesses a well-defined simultaneous position and momentum. From the standpoint of current quantum statistical mechanics, such a claim, as we should by now anticipate, entails metalogical self-referential inconsistency. This is why:

The Copenhagen results incorporate the limitative results of quantum indeterminacy. As we have seen, indeterminacy relations among conjugate variables do not reflect the technical limitations of laboratory physics at a

---

<sup>320</sup> A portion of this section is based on Bartlett (1988, p. 230).



particular point in time, limitations that might be superseded if more sensitive, sophisticated instrumentation can be developed, but they rather constrain, in principle, the range of what can identifyingly be referred to on the quantum level; they constrain what may both be measured and be meaningfully stated in the context of quantum theory. As we saw in the preceding section, the limitative nature of the Copenhagen results was vigorously opposed by a variety of theoretical physicists, including, as we've seen, not only Einstein, Podolsky, and Rosen, but also de Broglie, Jeffries, and Bohm. Of the arguments which were proposed at that time, perhaps David Bohm's is the only one which does not reduce quickly and simply to a dogged commitment without rational justification to traditional prejudices in favor of realism and complete physical determinism. Although it will not be to our purpose to go into the details of his view here, we may note that Bohm's rejection of the postulate of uncertainty never evolved into more than a hopeful sketch of an alternative quantum theory, one which received a skeptical response from most of his physicist colleagues and was eventually discarded even by him.<sup>321</sup> Nevertheless, in hindsight, Bohm's view is instructive as an example of what happens to quantum theory when it runs aground on the shoals of classical realist ontology.

Bohm has been the main defender of the hidden variable proposal.<sup>322</sup> From the standpoint of modern quantum theory, the hidden variable hypothesis is projective: It is metalogically self-undermining. Very briefly, the argument to demonstrate this may be stated as follows:

It can be shown that the uncertainty relations have the status of presuppositions in modern quantum mechanics. One way to do this is to show that from an operationally based statement of the uncertainty relations, the rest of quantum mechanics can be derived. This was proved by von Neumann in 1955.<sup>323</sup> However, it can also be shown that a denial of the postulate of uncertainty entails a denial of preconditions that must be satisfied in order for physical reference to specified dynamical variables—position/momentum, energy/time, or number/phase, etc.—to be possible. This is straightforward to establish:

---

<sup>321</sup> Among them, Heisenberg, Oppenheimer, Dirac, and Bethe expressed their strongest doubts concerning Bohm's proposal (in personal communications with Norwood Russell Hanson). See Hanson (1958, p. 174).

<sup>322</sup> The basis for Bohm's hidden variable proposal was laid earlier by de Broglie (1926) when de Broglie was still in his youth. He later abandoned his early hidden variable ideas; see also de Broglie (1964).

A more detailed analysis focusing on Bohm's hidden variable proposal, showing that it entails projective meaninglessness, may be found in Bartlett (1980, Section VII, pp. 161-167).

<sup>323</sup> Von Neumann (1955, Chapters IV, VI, especially pp. 323ff).

The algebraic analog of a statement simultaneously specifying precisely defined values for position and momentum itself is without meaning in quantum mechanics. This absence of meaning is due to conflict with the rules of formation and transformation employed in the formalism. But there is another, perhaps more interesting, reason for its meaninglessness.

As long as an alternative, comparably detailed and physically successful quantum theory is unavailable, the physical meaningfulness of a claim relating to quantum phenomena—e.g., relating to mutually interfering observables—will be understood in terms of prevailing quantum statistical theory. The uncertainty relations have the status of presuppositions, conceived of as rule-based constraints that govern the conceptual structure of the theory. The uncertainty relations are nothing more than the expression of a limitative postulate required in a calculus of operators.

A hidden variable theorist presumably wishes to refer to quantum phenomena as currently understood within the context of existing quantum theory; the proponent of hidden variables does not seek to discard the general established results of quantum theory, but wishes to construe the Copenhagen results, and perhaps extend them in a fashion consistent with those results, so as to satisfy the desires of realism. He wishes, specifically, to claim that mutually interfering observables “in reality” possess well-defined simultaneous values. Such a claim is clearly projective: The hidden variable theorist refers to a pair of observables that are essentially defined in a noncommuting sense, and in so doing explicitly denies a condition metalogically forced on the current quantum-theoretical understanding of interfering observables. The condition he denies is a precondition that must be satisfied in order for it to be possible for him, or anyone else, to refer meaningfully in the theoretical context in question to such observables. It is not that what the hidden variable theorist proposes is self-falsifying; rather, his claim is self-undermining in terms of its possible meaningfulness. The motivation to formulate a hidden variable theory of quantum mechanics, however, did not of course die out as a result of the post-mortem metalogical reasoning described here, but rather as a result of fundamental objections by quantum physicists and intractable difficulties in constructing the theory itself.

### **27.6 Quantum theory and the projection of “underlying reality”**

What we observe is not nature itself, but nature exposed to our method of questioning.

– Werner Heisenberg (1958, p. 58)

As we've seen already in connection with the EPR opposition to the Copenhagen results and also in connection with hidden variable proposals, the wish does not die out easily to assert, postulate, or at least hope for a physical reality that exceeds the referential capacity of frameworks of reference. When successive observations are made of the state of a quantum system, this wish lives on by insistently seeking to refer to what "must happen" "in between" those observations. This is an evident desire for framework-autonomous physical continuity; it expresses a resistance and rejection that quantum reality—and certainly that macroscopic reality—is inherently discontinuous.

In terms of a metalogical analysis of the preconditions of quantum reference, Heisenberg's reaction to this realist desire was insufficiently explicit, inadequately strong, and excessively restrained. For example, he wrote:

Quite generally there is no way of describing what happens between two consecutive observations.... [I]n quantum theory it would be a misuse of the language [to attempt to use language to describe anything pertaining to, e.g., an electron between observations] which ... cannot be justified....

The demand to "describe what happens" in the quantum-theoretical process between two successive observations is a contradiction *in adjecto*,<sup>324</sup> since the word "describe" refers to the use of the classical concepts, while these concepts cannot be applied in the space between observations; they can only be applied at the points of observation. (Heisenberg, 1958, pp. 48, 145)

Bohr, Heisenberg's teacher and mentor, considered things in much the same light. He attributed the issue to "the *difficulties* in talking about properties of atomic objects independent of the conditions of observation" (Bohr, 1958, p. 98, italics added). We recall that when he responded to the EPR position, he attributed the issue to an "essential *ambiguity*" in that position.

In fact what are involved are neither *difficulties*, nor an *ambiguity* in talking or thinking about this subject, nor an unjustifiable *misuse of language*, nor a "*contradiction in adjecto*," by which Heisenberg apparently meant a misapplication of classical concepts. It is rather a case of conceptually self-undermining reference on the level of possibility—self-undermining reference that results in meaninglessness.

This recognition is important, and were attention—instead perhaps of a

---

<sup>324</sup> I.e., a contradiction resulting from a noun and a modifying adjective which mutually exclude one another—e.g., a square circle.

largely willful blindness ({2.1})—to be given to this recognition, much of the energy expended in spinning alternative interpretations of quantum theory might be saved.

Milič Čapek (in Meyerson, 1985/1925, pp. xlvii, l) pointed to the reliance of realist physicists upon “the concept of permanent substratum,” which he associated with Piaget’s “concept of the object persisting in time [that] is formed in very early childhood.... Is it not then natural that such a concept, being a result of our cognitive adaptation to limited strata of reality, fails beyond its limits?” Čapek suggested that it is plausible that the notion of “persisting object” derived from childhood simply fails when applied later by adults when they seek to go beyond ordinary macroscopic objects with which they are familiar, to events on a quantum or cosmological scale. I mention Čapek’s remark because it contributes to a perspective that we generally lack: to see that the desire for continuously persisting reality that “exists independently of all possible means by which we may refer to it” is *conceptually primitive*, and certainly that desire may derive from a developmentally as well as an anthropologically primitive source. Čapek faulted the “inertia of our language and of our imagination” (Meyerson, 1985, p. l), but, again, it is not language and imagination that are responsible, but the difficult-to-restrain drive to trespass beyond the metalogical horizons of our frames of reference.

There is a natural human tendency to swing to the opposite extreme, and claim that “what cannot be observed does *not* exist,” a phrase von Weizsäcker employed in describing how the Copenhagen results are often misrepresented (von Weizsäcker, 1980/1971, p. 183, italics added). This contrary assertion, too, undermines itself by seeking to apply concepts beyond the horizons of their possible meaningfulness (see {21.5}).

Where then does this leave us? In the absence of the possibility of consistent inclusion in the mathematical formulation of quantum theory and the absence of the possibility of physical measurement, the physical question whether a quantum phenomenon “exists independently” is without possible meaning. To make this meta-level assertion of impossibility is not to assert a bare operationalist criterion of meaning, but it is to recognize that the nature of identifying reference in physics, and here in quantum physics, requires, as a precondition of the possibility of such physical reference, both the abstract coordinative structure of quantum theory and its correlation with a dynamic system formed by indissociable interrelations among observer, observational instrumentation, and—in metalogical consequence—the capacity in principle to observe and measure relevant quantum phenomena. It is only when these interrelations are enabled by such a dynamic system of reference that the meaningfulness of quantum concepts and statements made using them

becomes possible. “Nature itself,” to revise Heisenberg statement, has no possible meaning; rather, what we observe *is* nature.

### 27.7 Indeterminacy and uncertainty

In quantum mechanics, we are not dealing with an arbitrary renunciation of a more detailed analysis of atomic phenomena, but with a recognition that such an analysis is *in principle* excluded.

– Niels Bohr (1958, p. 62)

Having arrived in the preceding pages at a de-projected understanding of the EPR position, we recognize that the desire for an ontology advocating quantum-theoretical realism is metalogically self-undermining and without meaning. Heisenberg’s principle of uncertainty now needs to be understood in this light.

Heisenberg was apparently and justifiably ambivalent when he chose the most appropriate name for his principle. In his earliest publication in which he developed the principle (Heisenberg, 1927), he most frequently employed the German word ‘*Ungenauigkeit*’, which has the meaning of ‘the state of being inexact, imprecise, or indeterminate’. The German word ‘*Unsicherheit*’, in contrast, means ‘uncertainty’, a term he used in his paper only twice. In the article’s Addendum, written after he’d completed the paper, Heisenberg introduced the term ‘*Unsicherheitsrelation*’, meaning ‘uncertainty relation’. In subsequent publications of his work translated into English, ‘uncertainty’ rather than ‘indeterminacy’ is most often used; as a result, the name ‘uncertainty principle’ has become the most familiar.

The use of the term ‘uncertainty’ has had the effect of implying a putative uncertainty relating to the simultaneous quantum measurements of conjugate properties, and of suggesting a physicist’s resulting purported *lack of certainty* in making predictions concerning such properties. Similar both to Einstein’s regretted choice in naming his special and general theories, and to Heisenberg’s disappointment in labeling the Copenhagen results an ‘interpretation’, in this instance Heisenberg was, as I understand his intention, inclined to accept a name for his principle which misplaced and detracted from his desired emphasis: While his intention was clearly to describe the *indeterminability in principle* that conjugate properties have simultaneous precisely defined values, the name ‘the uncertainty principle’ instead pointed to uncertainty of simultaneous measurements of conjugate properties and resulting uncertainty

of the knowledge of the observer. The difference relates, on the one hand, to the nature of quantum reality—that certain of the properties of quantum phenomena are fundamentally *inexact* (*ungenau*) or *indeterminate*—and, on the other hand, to a *lack of certainty* (*Unsicherheit*) of the measurements of quantum properties and of the physicist’s resulting knowledge. It is clear that a description of a physicist’s degree of uncertainty is not equivalent to a description of the indeterminate nature of a quantum phenomenon; these are two different kinds of description, with two different emphases. To express this difference metaphorically: The quantum phenomena Heisenberg sought to describe do not, when conjugate properties are in view, have sharply etched, *exact* edges; the fact that such phenomena are, in important ways, inexact or indeterminate is made evident in terms of the inability in principle of quantum physics to predict the outcome of many quantum measurements with exactitude. In these two senses, “indeterminacy” and “uncertainty” refer to closely related aspects of the same thing; indeed, in this context they are complementary terms, but this is “complementarity” of a separate kind than is expressed by the principle of complementarity.

These two closely related meanings of ‘indeterminacy’ and ‘uncertainty’ are two aspects of the same realization. There is no knowledge which quantum physics “falls short of having” in the sense of “uncertainty,” for the fundamental indeterminacy of quantum phenomena provides all the knowledge that is possible of them. Indeterminacy, randomness, and probability distributions, which so rankled Einstein, are not signs of a *deficiency* of knowledge, which the choice of the term ‘uncertainty’ suggests. The denial of this fact—established both by the Copenhagen results and by their understanding in terms of the metalogic of reference—leads to projective meaninglessness. The claim that would express such a denial cannot, in principle, meaningfully be made within the framework of Copenhagen quantum theory in a manner so that preconditions of reference that would allow for the possibility of that denial can be satisfied. Once again, as we have seen in other cases in which attempts are made to transgress beyond the horizons of frameworks of reference, we encounter a putatively meaningful claim that undermines its own possible meaning. The purportedly meaningful belief that indeterminacy or uncertainty, randomness, and the stochastic predictions of quantum theory reflect deficient knowledge is a claim precisely of this kind.<sup>325</sup>

---

<sup>325</sup> Heisenberg pointed to the tendency of some quantum physicists to make meaningless statements in connection with the uncertainty relation between measurements of position and momentum. He wrote: “This uncertainty relation [i.e.,  $\Delta x \Delta p_x \geq h$ ] specifies the limits within which the particle picture can be applied. Any use of the words ‘position’ and ‘velocity’ with an accuracy exceeding that given by [this] equation ... is just as *meaningless* as the use of words

The Copenhagen results, as we may begin to see them through the lens of the metalogic of reference, do not *deny*, as is commonly thought, the existence of a fundamental quantum reality independent of its observability (a denial which, as we've seen in previous chapters, is a meaningless, projective claim), nor do they *affirm* such an autonomous quantum reality (which is equally projective). They instead show that what we understand, by virtue of those results, *is* quantum reality. It is a mistake to consider these results as offering *merely* an abstract means that make it possible formally to express and then to calculate experimental results; the Copenhagen results provide us with *a new way of conceptualizing and of understanding the nature of physical reality on the quantum level*. They also establish limitations on what we may, with possible meaning, claim about that physical reality.

Einstein, Podolsky, Rosen, and other physicists who opposed the Copenhagen results preferred to believe that indeterminacy and the randomness associated with it which are evident on the quantum level, coupled with the need to make recourse to probability functions, reflect human ignorance concerning underlying, framework-independent reality. They were unwilling to accept that quantum physics should be—from their viewpoint—condemned to an unsatisfactory state of permanent “uncertainty.” Bohr disagreed entirely, claiming that quantum *indeterminacy is a fundamental property of micro-physical reality*: Properties of quantum phenomena that are indeterminate are not partial properties that call out to be filled by data from further, more exact observations. As we have seen, a strongly compelling justification for Bohr's disagreement is provided by the metalogic of reference. We shall explore the implications of this justification as we examine Bohr's concept of complementarity.

---

whose sense is not defined” (Heisenberg, 1930, p. 15, italics added). He then added a footnote to this passage:

[O]ne should particularly remember that the human language permits the construction of sentences which do not involve any consequences and which therefore *have no content at all*—in spite of the fact that these sentences produce some kind of picture in our imagination; e.g., the statement that besides our world there exists another world, with which any connection is impossible in principle, does not lead to any experimental consequence, but does produce a kind of picture in the mind. Obviously such a statement can neither be proved nor disproved. One should be especially careful in using the words ‘reality’, ‘actually’, etc., since these words very often lead to statements of the type just mentioned. (Heisenberg, 1930, p. 15n, italics added)

### 27.8 Complementarity: “*Contraria sunt complementa*”<sup>326</sup>

In previous chapters, two abstract senses of complementarity have been described. In {7.3.6} note was made of inconsistency-tolerant and inconsistency-asserting approaches to mathematical logic from the standpoint of which conflicting propositions may both be countenanced; in this sense, such formal systems may be considered to express a logical variety of “complementarity.”

In {10.5} a different sense of ‘complementarity’ was introduced in connection with the concept of “perspectives.” Two reference frames establish two different perspectives when two sets of objects of reference, a set being given in each reference frame, are considered to be either “the same,” or “complementary” in the following sense: The concept of dimensional incompatibility of frameworks of reference was introduced; that concept applies, to give an example, in the case when one framework of reference permits reference to 2-dimensional objects, in which the language of “surfaces” has meaning, and a second framework allows for reference to 3-dimensional objects, in which the language of “volumes” can be used. The 3-dimensional concept of volumes does not retain its meaning in the 2-dimensional framework. In this sense, the two frameworks are said to be “dimensionally incompatible.” When two frameworks of reference are dimensionally incompatible, they may be considered to be “complementary” when, from the standpoint of a third—a meta-level—framework, the sets of objects that may be identified in one framework qualifies as the same set of objects identifiable in the other. Such frameworks of reference may be considered to express a framework-relative variety of “complementarity.” Fundamentally presupposed in applications of such a concept of complementarity is, of course, the conceptually problematic notion of “the sameness of objects of reference,” briefly discussed in the first two sections of the preceding chapter. This the problem of correlating phenomena from the standpoint of different reference frames. This problem becomes particularly manifest when reference frames are dimensionally incompatible, a situation which we encounter in a very explicit way in quantum theory. It is the concept of the complementarity of frameworks of reference as it applies to quantum theory that will concern us in this section.

As noted in {10.5}, the unifying presumption that reference from the standpoint of two dimensionally incompatible frameworks involves one and the same “underlying” object of reference is a presumption that cannot be made without qualification, as when two mutually exclusive laboratory setups

---

<sup>326</sup> “Opposites are complementary,” the motto Bohr placed on the coat-of-arms that he designed for himself. (Cf. Wheeler, 1985.)



alternatively make it possible to refer to light as a particle phenomenon or as a wave phenomenon. Heisenberg stated this presumption boldly and seemingly without concern: “The problem of quantum theory centers on the fact that the particle picture and the wave picture are merely two different aspects of *one and the same physical reality*” (Heisenberg, 1930, p. 177, italics added). Bohr’s principle of complementarity functions as a *meta-level assertion* (i.e., formulated from the standpoint of the meta-level framework noted in the previous paragraph); that assertion supplies the needed qualification: that both quantum phenomena—particle and wave—are to be *correlated*, and that, in terms of that correlation, are to be understood as mutually exclusive aspects of one and the same quantum phenomenon when observed under two mutually exclusive experimental arrangements. As Bohr expressed this:

[I]t is only the mutual exclusion of any two experimental procedures, permitting the unambiguous definition of complementary physical quantities, which provides room for new physical laws, the coexistence of which might at first sight appear irreconcilable with the basic principles of science. It is just this entirely new situation as regards the description of physical phenomena, that the notion of *complementarity* aims at characterizing. (Bohr, 1935, p. 700)

The study of quantum phenomena that are complementary requires mutually exclusive experimental arrangements—which is to say, observational reference frames that are, in our terms, dimensionally incompatible. The information obtainable from complementary reference frames is *complete information* about the properties of conjugate quantum phenomena: The Copenhagen results demonstrated this, while the methodology of the present study shows that a denial of this claim to completeness is metalogically self-undermining and therefore devoid of meaning. As Bohr expressed this conclusion:

However great the contrasts exhibited by atomic phenomena under different experimental conditions, such phenomena must be termed complementary in the sense that each is well defined and that *together they exhaust all definable knowledge about the objects concerned*. The quantum-mechanical formalism, the sole aim of which is the comprehension of observations obtained under experimental conditions described by simple physical concepts, gives just such an *exhaustive*

*complementary account* of a very large domain of experience....

In order to characterize the relation between phenomena observed under different experimental conditions, one has introduced the term complementarity to emphasize that such phenomena together *exhaust all definable information about the atomic objects*. Far from containing any arbitrary renunciation of customary physical explanation, the notion of complementarity refers directly to our position as observers in a domain of experience where unambiguous application of the concepts used in the description of phenomena depends essentially on the conditions of observation. (Bohr, 1958, pp. 90, 99)

This is a clear and explicit statement of framework relativity as it applies to quantum phenomena. An acceptance of such framework relativity brings with it, as seen earlier in this chapter, a need to renounce the EPR notion that quantum phenomena exist in some putatively meaningful way independently of their observability. This renunciation has been difficult and painful for some physicists to accept, and impossible for others like Einstein. An added burden has been the need to renounce a bias in favor of physical continuity that had been unquestioning accepted by classical physics: that between successive observations of the state of a quantum system, continuously existing physical reality must nonetheless “be there” despite quantum randomness, discontinuity, and the essential need to accept that the predictability of quantum events is governed by probability functions. In addition to these renunciations, due to the inescapable framework relativity of the dynamic system we’ve described, it has also been necessary to relinquish the traditional notion of causality. From Bohr’s point of view, as I have quoted earlier in this chapter,

... the finite interaction between object and measuring agencies conditioned by the very existence of the quantum of action entails—because of the impossibility of controlling the reaction of the object on the measuring instruments if these are to serve their purpose—the necessity of a final renunciation of the classical ideal of causality and a radical revision of our attitude towards the problem of physical reality. (Bohr, 1935, p. 697)

Supporting this observation by Bohr from the standpoint of the metalogic of reference, we've shown in {23} that the classical idea of causality is untenable because it is inherently projective.

A full acceptance of quantum complementarity is seldom described or appreciated with explicit clarity: It requires that one understand and acknowledge, without reservation and with a gradually learned sense of emotional and intellectual comfort, that physical reality—in this chapter speaking only of the quantum domain—is “no more” than the quantum phenomena to which physical identifying reference is possible relative to frameworks that assure that possibility. This is all that we have and can in principle have; this is all we can meaningfully ask for, and it is also all that is possible. As we've seen, that framework is a dynamic structure, consisting of a theoretically abstract coordinative mathematical formulation in terms of which indissociable interrelations are necessary among observer, instrumentation, and the capacity to observe and measure the small-scale phenomena of interest. Bohr's principle of complementarity is most commonly understood, by both quantum physicists and philosophers of physics, as “a resignation in the face of incomprehensible empirical difficulties in the measurement process” (von Weizsäcker, 1980/1971, p. 399), but to believe that “resignation” is required, is to miss the point. One cannot resign oneself to the absence of something whose autonomous existence *or* nonexistence is without possible meaning.

### **27.9 The inseparability of the observer and the observed**

One of the main sources of conceptual difficulty for which quantum theory is responsible concerns the putatively meaningful separation of the observer together with the observer's instrumentation, from the quantum phenomena that are observed. Earlier in this chapter I briefly considered what has come largely to be accepted to be the presumably meaningful notion that quantum measurements “perturb” what is measured. We have only touched upon the question whether it is possible in principle for relevant referential preconditions to be satisfied which are necessary in order possibly to assert that a “change” in the observed quantum phenomena has “come about following quantum measurement.” Fundamental to this question is the potential meaningful separability or distinguishability of, on the one hand, the observer and the instrumentation employed by the observer, and, on the other hand, the quantum phenomena measured. Certainly, at least on the surface, to claim that “the observer + measuring apparatus ‘perturbs’ the quantum phenomena measured” would seem to imply a relationship of “interaction” between the observer + measuring apparatus and the quantum objects measured, and hence

their separability—or at the very least, their distinguishability in some meaningful sense—since this is assumed by the relational nature of the claim. We must therefore ask whether such a relational understanding can, in principle, make sense here.

Among quantum physicists, Bohr was especially aware both of the physical limitations involved in this question and of the need for clarity and precision from an epistemological standpoint. He emphasized the “*coupling* between phenomena and their observation in which the finite magnitude of the quantum of action prevents altogether a sharp distinction being made between a phenomenon and the agency by which it is observed” (Bohr, 1934, pp. 10-11, italics added). With respect to the principle of complementarity, he stressed that the principle “implies the impossibility of any sharp separation between the behaviour of atomic objects and the interaction with the measuring instruments that serve to define the conditions under which the phenomena appear” (Bohr, 1996, p. 210).

If it is taken for granted that a relation of “agency-based perturbation” occurs when quantum phenomena are measured by an observer, then already built into this notion is the putatively meaningful claim that the observer exercises “agency” and that an “interaction” takes place between the measuring apparatus and the quantum phenomena measured. Bohr made use of such notions, while at the same time cautioning that no “sharp distinction” or “sharp separation” can be made between the “agent” and the “phenomena.” These terms and the notions they expressed do not, as we shall see, serve Bohr’s purpose: The routine meaning of the notion of “interaction” is “action between two things,” while “agency” is commonly understood as “activity undertaken by an agent directed at something other than the agent.” We are forced to ask, How are the observer, including the observer’s instrumentation, and the observed quantum phenomena to be understood if there is no sharp distinction or separation in the context of quantum measurements that can meaningfully be made between them? If it is not a “sharp separation,” is it a “fuzzy separation,” or is it entirely inappropriate, and perhaps metalogically self-undermining, to call or think of this as a “separation” at all? Certain of the results reached in previous chapters can provide the needed clarification.

Earlier in this chapter, we recognized that the Copenhagen results entail that it is without meaning to ascribe properties to an isolated quantum phenomenon because it is in principle impossible to separate such a phenomenon from its functional relation defined by the dynamic system consisting of observer–measurement apparatus–quantum object. We’ve called this ‘the framework relativity of quantum phenomena’.

In this study, framework relativity has been characterized ({10.2–10.3})

by both the self-enclosure of the referential field associated with a given frame of reference, and by the embedment in the identity of an object of reference of its constitutive structure, the preconditions necessary for its identifiability. In these terms, the identity of an object of reference is an instantiation or realization of the referential field of its presupposed frame of reference. We also recall ({25.2}) the following properties of referential fields which are relevant to the question of the nature of a “separation” between quantum phenomena and observers supported by their instrumentation:

- a referential field defines a functional interdependency between the identities of a set of possible objects of reference within the field and an associated reference frame
- The embedded constitutive structure of objects of reference in the above set determines their in-principle identifiability as a function of the field
- Specification of an object of reference involves a dynamic inter-relationship between the field and the object, i.e., possible properties of an object of reference are functionally defined by its referential field
- A referential field considered without explicit reference to objects in that field is only a *potential* basis for a set of in-principle identifiable objects
- A referential field satisfying the above conditions is such that coordinate positions in the field can be associated with a potential observer.

We have recognized that the field concept in physics is fundamentally a modal concept: The field effect in physics, as well as the central meaning of the concept of referential fields, are essentially descriptions of a formally envisaged range of possibilities under a specified set of constraints. Understood in these terms, we have come to see ({14.2}) that a referential field possesses “reactive” boundaries which make evident the field’s metalogical horizon; if these boundaries are violated possible reference and meaning are undermined.

These results may now be applied to the question how we are to understand quantum phenomena in the framework-relative terms of this study. Both the mathematical formulation of quantum theory and the physical phenomena

to which it applies have made explicitly evident not only the need to recognize that both mathematical and physical identifying reference to them is, in principle, conditioned by the frameworks in terms of which such reference is possible, but also the need to change the way in which we think about such framework relativity. In quantum theory, it is no longer possible to make any meaningful distinction between quantum phenomena and the reference frames in terms of which they are identified, something we are able to do in concrete terms, for example, when associating the lines and information on a road map with their physical highway counterparts. When quantum phenomena are involved, there does not exist a separable “road map” or “distinguishable reference frame”; instead, the referential field that establishes the modal basis for a range of possibilities under a particular set of constraints comprises a dynamic system with a formalized mathematical expression, a dynamic system in terms of which we cannot, in principle, apply the traditional notions of “interaction,” “agency,” and “measurement-based perturbation.” These notions trace their origins back to a set of conceptually primitive beliefs, many of which we’ve shown in earlier chapters to be metalogically projective, and some of these, which are employed in conceptualizing quantum theory, can have no possible meaningful application in the quantum domain. The wave equation expresses in mathematical form the referential field defined by a given dynamic quantum system. It can, in principle, make no sense to ask questions relating to the state of a particular quantum object of reference independently of the referential field which makes reference to that object possible. Nor is it possible to attach any sense to efforts that would seek to distinguish the role of the observer, the observer’s instrumentation, and the quantum system in view, which the notions of “interaction,” “agency,” and “measurement-based perturbation” presuppose.

Bohr was right when he repeatedly used phrases such as ‘implies the impossibility of ...’ in connection with the principle of uncertainty and in connection with what I have called ‘the framework relativity of quantum phenomena’. The *impossibility* that he had in mind is a meaning of ‘impossibility’ that combines both physical and mathematical impossibility. His description of quantum measurement in terms of “interaction” that “perturbs” a “pre-existing quantum state”—a “pre-existing state” which, if one accepts the very notion of measurement-based perturbation, cannot possibly be observed or measured, and whose existence can only be posited as a self-undermining projection—was motivated by Bohr’s most careful of intentions, but fell short for lack of epistemological rigor.

As we have seen in a seemingly almost endless proliferation in a wide variety of contexts, the temptation and the drive to trespass beyond metalogi-

cal horizons are difficult to place under rational control. In the final sections of this chapter, we shall attempt to describe in clear and non-projective terms how it is possible to understand the framework-relative nature of quantum phenomena in a manner freed from the relational subject-object, agency-based bias that has for so long influenced even the thought of leading quantum physicists.

### 27.10 The “collapse” of the Schrödinger wave function

The Copenhagen results, which have exercised such a strong influence on the subsequent development of quantum theory, have inspired a considerable amount of metalogically projective thinking. Much of this projective thinking has been expressed in the form of “interpretations” to which the Copenhagen results have been subjected. In the contest of interpreting voices, one of the often-discussed and often-argued subjects related to the Copenhagen results concerns the so-called ‘collapse’ of the Schrödinger wave function. Since the controversies that have ensued over the “collapse” continue to this day, and since the approach of this study may shed some clarifying light on the issue, I include a few words about it.

In abbreviated and general terms, the Schrödinger wave function is an abstract mathematical representation of the theoretical range of possibilities that can be associated with the state of a quantum system. From the standpoint of the Copenhagen physicists, the Schrödinger wave equation comprises a complete description of a quantum system; in fact, the equation has been considered to formulate the most complete description that *can* be given of a quantum system. The equation was named after Erwin Schrödinger (1887–1961); the equation was published in 1926, and led in 1933 to Schrödinger’s Nobel Prize. Although the equation can be used to calculate predictions of measurement results of important properties of a quantum system, there has been considerable controversy over what the status—epistemological or ontological—of the equation actually is; this controversy has left unanswered important questions concerning interrelationships among the abstractly formulated wave function, the results of quantum measurements, and, as we should by now have come to expect, “underlying physical reality.”

Individual measurements of a quantum system always provide information that the system is in some particular (the measured) state. The wave function, however, describes how a quantum system theoretically evolves; this evolution is mathematically expressed as a linear superposition of many

possible values.<sup>327</sup> Since an individual measurement of a quantum system always finds the system in some particular measured state, physicists who adhere to the measurement-based perturbation view have wished to know how the superposition of many possible values described by the wave equation “becomes” an individual measured value. The wish to know this has led to the widely discussed “measurement problem” in quantum theory. Some of these physicists have embraced the notion that Schrödinger’s equation describes an “underlying reality” which then takes on a particular value “as the result” of the “perturbation” brought about through the “act” of measurement.<sup>328</sup>

Readers who are by now familiar with the methodology of the present study will not require further explanation to show that the foregoing notion is, in principle, devoid of meaning because it is metalogically projective in multiple ways. (Quotation marks placed around problematic words in the preceding paragraph serve as “red flags” of the projections involved.) Instead, let us consider how Bohr and Heisenberg sought to avoid such a confused and self-undermining view.

Bohr considered Schrödinger’s wave equation to be an important conceptualization of the mathematics of quantum theory, but did not claim that the equation could be associated with a corresponding physical structure in nature. In his estimation, the wave equation is purely a mathematical tool that describes a quantum system. By 1927, Heisenberg had apparently agreed with this understanding.<sup>329</sup> Their shared position was deliberately restrained and made no more than this minimal claim.

However, later in Heisenberg’s life, perhaps due in large part to the increasingly contested status of the wave function’s “collapse,” Heisenberg chose to describe the “reduction” of the wave equation, when a quantum system is measured, in terms of a shift from “possibility” (alternatively, “*potentia*,” “probability,” or “objective tendency”) to “actuality”:

[T]he transition from the “possible” to the “actual” takes place during the act of observation. If we want to describe what happens in an atomic event, we have to realize that the

---

<sup>327</sup> Superposition in classical wave mechanics means that two waves that propagate in the same space will combine to form a wave with a net amplitude equal to the sum of the amplitudes of the two individual waves; the summation of the individual amplitudes is their “superposition,” which is “linear” when the amplitudes of the waves add together in this way. Similarly in quantum theory, two or more quantum states can be added together to yield their superposition, which is linear when their summation produces another valid quantum state.

<sup>328</sup> With certain qualifications and variations, examples include Penrose (1999/1989, pp. 475-481) and Ghirardi, Rimini, and Weber (1986).

<sup>329</sup> Heelan (2017, p. 114).



word “happens” can apply only to the observation, not to the state of affairs between two observations.... The observer has ... only the function of registering decisions, i.e., processes in space and time, and it does not matter whether the observer is an apparatus or a human being; but the registration, i.e., the transition from the “possible” to the “actual,” is absolutely necessary here and cannot be omitted from the interpretation of quantum theory. (Heisenberg, 1958, pp. 55, 137)

Heisenberg’s description, however, became less clear, laden with added epistemological problems, when he sought to explain what this “transition” from “possibility” to “actuality” means:

[T]he transition is completed from the possible to the actual [when] ... [t]he probability function, which covered a wide range of possibilities, is suddenly reduced to a much narrower range by the fact that experiment has led to a definite result, that actually a certain event has happened. *In the formalism* this reduction requires that the so-called interference of probabilities, which is the most characteristic phenomenon of quantum theory, is destroyed by the partly undefinable and irreversible interactions of the system with the measuring apparatus and the rest of the world. (Heisenberg, 1958, p. 142, italics added)

The loss of clarity occurs in the last sentence: On the one hand, Heisenberg appears to limit the “transition” to the context of the mathematics employed, that is, to the formalism, but on the other hand, he succumbs to the notion of “partly undefinable and irreversible interactions of the system with the measuring apparatus and the rest of the world,” a notion which, as we’ve seen, is inherently projective. Sometimes a thinker will continue to work on a problem beyond its need for further work, losing sight of a more adequate solution he or she left behind. This is, to my mind, what happened to Heisenberg.<sup>330</sup>

---

<sup>330</sup> It is further exemplified when Heisenberg stretched further in attempting to explain whether his *potentia* allegedly involved in the transition really exist: ‘*Potentia*’ refers to “something in the middle between the idea of an event and the actualization of the event, *a strange kind of physical reality* in the middle between possibility and actuality” (Heisenberg, 1958, p. 41, italics added). “A strange kind of physical reality” that lies “in the middle between an event and the actualization of the event” is objectionably vague and uninformative, but more importantly, such a notion is metalogically projective. —Once again, we see in even a careful thinker the urge to overstep the metalogical horizon established by his own thought.

Nearly a century has passed since Bohr and Heisenberg were satisfied with a minimalist recognition that the wave equation is a formal tool of quantum theory, and were satisfied not to attempt to go beyond this. This minimalist understanding may well be a complete understanding, since to claim more than this leads to impassables from which it has been difficult for later physicists to extricate themselves—largely, I suggest, because of an inability to recognize the metalogically projective nature of the questions they raise.

The minimal position that was in the early years adopted first by Bohr and then also by Heisenberg is a position that has continued to be voiced, now and again, but with mixed clarity, in the midst of the impassioned controversy over the “collapse” of the wave function. I quote one example: “Various conceptual difficulties disappear when one realizes that collapse is something which takes place in the theoretical physicist’s notebook and *not* in the experimental physicist’s laboratory” (Griffiths, 2002, p. 9, italics added). In discussing ways to avoid the conceptual difficulties attendant to talk about “wave function collapse,” Griffiths recommends

... to think of wave function collapse not as a *physical effect* produced by the measuring apparatus, but as a *mathematical procedure* for calculating statistical correlations.... “[C]ollapse” is something which takes place in the theorist’s notebook, *rather than* the experimentalist’s laboratory. (Griffiths, 2002, p. 247, last italics added)

I’ve characterized these more recent claims as involving “mixed clarity,” for they, too, yield to the temptation to attempt to say more than they can, in principle, meaningfully claim: It is not that the wave function either *does*, or *does not*, “collapse” *only* in the physicist’s laboratory, for to claim either of these is to seek to make a claim whose referential preconditions cannot, in principle, be satisfied. We can, however, make the meaningful claim that Schrödinger’s wave equation comprises an important and useful mathematical formulation, and *in that formal context*, the “reduction” or “collapse” of the wave function, when individual measurements of a quantum system are made, can make sense.

### **27.11 The non-relational ontology of quantum phenomena**

An interaction described purely quantum mechanically takes the form of an internal dynamics of a composite object

consisting of all the interacting objects; the original object has become “submerged” in this totality. The original object itself is measured only when an irreversible event occurs in the objects with which it interacts, and which we call the instrument of measurement...

[T]he concept of interaction corresponds to the approximative way of talking on which all physics rests: we speak of separate objects or separate alternatives, knowing that they do not exist in a strict sense, and *we correct this mistake by describing them as interacting objects.*

– C. F. von Weizsäcker (1980/1971, pp. 395, 215, italics added)

In {25}, a non-relational understanding of reference was described in which the concept of reference was developed that departs radically from its traditional, commonsensical, subject-object, referrer-referent, activity-based conception. {23} and {24} showed that conventional causal claims and claims that involve a putative self in the role of agent are essentially projective, and cannot for that reason, in principle, be meaningful. The non-relational understanding of reference that has been developed makes it possible to avoid agency-based causal claims about the nature of reference. In their place, the concept of referential fields provides an understanding of the functionally interdependent nature of the organization of integrated totalities that form dynamic systems of reference.

It is important to remind the reader what is usually meant by the term ‘relation’. As noted in {5.1}, the word is most commonly used to claim that there is a connection or a contrast between relata that are differentiable—that is, things whose properties are such that the things possessing those properties can be distinguished from one another. I quoted Locke who had this in mind when he wrote: “The nature ... of Relation consists in the referring, or comparing two things, one to another” (Locke, 1690, p. 151). When in the present study I have used the term ‘non-relational’, it is to de-emphasize, and ideally to avoid entirely, the notion that it is putatively meaningful to distinguish the relata presumed to be involved.

A non-relational understanding of reference applies directly to the framework-relative realization by quantum theory that small-scale phenomena are identifiable only in terms of the dynamic system formed, in Bohr’s words, by the “*coupling* between phenomena and their observation” (italics added). As physicist Henry Stapp sought to express this:

[A]n elementary particle is not an independently existing, un-analyzable entity. It *is*, in essence, a set of relationships that reach outward to other things.... At the atomic and elementary-particle level, the idea of independent entities dissolves; the most elementary things have *meaning* only in terms of their effects on other things.... [O]ne finds in the realm of elementary-particle physics ... a web structure: Analysis never yields an ultimate set of unanalyzable basic entities or qualities. The smallest elements always reach out to other things and find their *meaning* and ground of being in these other things.... (Stapp, 1971, pp. 1310, 1314, 1319, italics added).

When we reflect on these claims, they may be thought to point, as I interpret Stapp's basic intentions, to a non-relational recognition of how the existence of elementary particles is to be understood—despite the obvious fact that in expressing this he made recourse to the habitual language and commonsense notions of “relationships” and “effects on other things.”<sup>331</sup>

In terms of referential fields, the identity of a quantum phenomenon *is* a functional interdependency. In relation to the conclusions reached in {18}, the ontology of quantum phenomena can be correspondingly understood: As was expressed there, for something to *be* is a function of the coordinative relation of object of reference in relation to an appropriate reference frame in terms of which it can be identified, whether vaguely or precisely, or according to rule. If one wishes to know the ontological status of a class of objects, one must look to the constitutive structure of the framework of reference in terms of which those objects are identifiable. Ingredient in any object of reference is that object's constitutive structure: i.e., the structural/systemic presuppositions of the reference frame in terms of which it is identifiable. The object, its very identity, is indissolubly linked to the reference frame(s) that provide the basis for its identifiability. The nature of that link is informative: It tells in what sense the object *is*.

We have seen in the present chapter how quantum theory has largely come to be viewed in terms of the “interaction” between measuring instru-

---

<sup>331</sup> I pass over details relating to Stapp's “S-matrix [scattering matrix] interpretation” of quantum theory and quote from Stapp's paper only because he appeared to point in the “non-relational” direction (although Stapp might not agree that was his intention). Beyond this, his paper represents a good example of the extreme conceptual difficulties that are encountered when attempting to maintain a projective distinction between the observer and the quantum phenomena measured, while at the same time recognizing that they form what I've called a ‘dynamically interdependent system’.

ments, the active agency of observers, and the quantum phenomena studied. We have also seen how an attempt to understand quantum theory in these terms is metalogically self-undermining. Bohr has claimed that “the ‘individuality’ of atomic processes, as conditioned by the quantum of action, will be frustrated by the unavoidable interaction between the atomic objects concerned and the measuring instruments indispensable for that purpose” (Bohr 1958, p. 19), while von Weizsäcker has acknowledged that it is a *mistake* to describe such objects as “interacting”: “we speak of separate objects or separate alternatives, knowing that they do not exist in a strict sense” (von Weizsäcker, 1980/1971, p. 215, italics added).

From the standpoint of the metalogic of reference, the identity of a quantum phenomenon *is* a functional interdependency: note that *functional interdependency is not “interaction.”* The sense in which a quantum object exists *is* as an object of reference indissolubly linked to the reference frame(s) that provide the basis for its identifiability. Ingredient in the quantum phenomenon’s very identity are the structural/systemic presuppositions of the reference frame(s) in terms of which it is identifiable. To speak of “interaction,” “measurement-based perturbation,” and the “active agency” of observers is not only a “mistake,” it is, as this chapter has made clear, conceptually incoherent because it undermines itself on a metalogical level. When understood through the lens of the metalogic of reference, Bohr’s statement can be viewed in a newer and clearer light, one which Bohr likely could not foresee: He stated that quantum theory (specifically, the discovery of the quantum of action)

... has thrown new light on the very foundation of the description of nature and revealed hitherto unnoticed presuppositions to the rational use of the concepts on which the communication of experience rests. In quantum physics, as we have seen, an account of the functioning of the measuring instruments is indispensable to the definition of phenomena....  
(Bohr, 1958, p. 91)

In the context of the present study, we recognize that these “unnoticed presuppositions” are not simply those of classical physics, but are suppositions which have been made in the absence of clear epistemologically focused referential analysis, suppositions which require fundamental, de-projective revision. When we have done this, we see more clearly how “an account of the functioning of measuring instruments is indispensable to the definition of phenomena,” but—more precisely—we see how this “indispensability” has clear-cut ontological consequences with respect to the quantum phenomena

investigated. More than this, a de-projective analysis serves to define fundamentally significant and inescapable horizons of possibility and meaning of the central concepts and claims to knowledge of quantum theory.

### 27.12 The reality of quantum discontinuity

[T]he element of discontinuity ... is found everywhere in atomic physics....

– Werner Heisenberg (1958, p. 143)

That one should treat given objects as noncontinuous is, I think, the fundamental truth of quantum theory. Quantum theory has eliminated the trouble created by continuity in classical physics ... and it has run into troubles of its own

where it attempted to build “true” continuity into its frame; i.e., in the field theory of interactions.

– C. F. von Weizsäcker (1980/1971, pp. 213-214)

There is no purely logical reason to believe that the evidenced discontinuity of quantum phenomena is somehow unique to the domain of microphysics, and to believe, in contrast to the quantum domain, that macrophysical reality is exempt from much the same kind of discontinuity. Physicists who, like Einstein, preferred to believe that observed quantum phenomena are manifestations of an underlying, independently existing reality, have similarly not wanted to accept discontinuity on the quantum level as ultimate. They have, as we’ve seen, wished for more than this; they have wished, as Heisenberg described their desire, that the observed discontinuity of quantum phenomena *leaves out*

... what actually happens independently of or between the observations. *But something must happen*, this [they] cannot doubt; this something need not be described in terms of electrons or waves or light quanta, but unless it is described somehow the task of physics is not completed. It cannot be admitted that it refers to the act of observation only. The [traditional realist] physicist must postulate in his science that

he is studying a world which he himself has not made and which would be present, essentially unchanged, if he were not there. (Heisenberg, 1958, p. 144)

The same metalogically self-undermining wishful thinking of course has, as one should expect, been expressed in connection with the randomness of quantum events and the unavoidability of recourse to the description of quantum systems in terms of probability functions.

Projectively realist beliefs have not, as we've seen in earlier chapters, been confined to quantum physics; they permeate many of the projective notions of common sense and of traditional philosophy. It is important that we be reminded of related projections we have previously identified.<sup>332</sup> We saw how habitual projective ways of thinking can lead to the putatively meaningful belief in an "independently existing physical space-time continuum" in which physical objects are "located." Once these purportedly meaningful notions have been de-projected, we saw that temporal and spatial continuity *mean* the expectation, whether as a result of habituation or through the application of rules, that forms the basis for the rational and meaningful concept that a given set of perspectives may, in principle, be continued in time, and may, again in principle, be further extended in space. But what can in many instances be a reasonable *expectation* is not at all what realist physicists and philosophers would wish; instead, they prefer to believe that such a set of perspectives is "in reality" a manifestation or expression of a putatively meaningful "independently existing temporal or spatial continuum"—and this, as we have found, entails a metalogically self-undermining projection of temporal or spatial continuity.

We have previously recognized that sensory experience is given in discontinuous perspectival form, and that in many cases perspectives from the standpoint of one framework of reference may be considered complementary in relation to those given in a different, dimensionally incompatible reference frame ({10.5, 21.1, 21.3, 27.8}). Quantum theory has, in parallel, come to recognize both the discontinuity of many observed quantum phenomena, as well as the complementarity of quantum measurements made by means of incompatible experimental arrangements. It should come as no surprise that when quantum physicists find a basis in the observed behavior of quantum phenomena which permits them to formulate principles of regularity that they should wish to infer the continuity of the objects to which they refer, despite discontinuities of their observations, and despite discontinuous properties of those phenomena. Continuity that "fills in the blanks" between periods of

---

<sup>332</sup> See, in particular, the discussion of the notion of and belief in continuity, {21.4} and {22.5}.

discontinuous observation, or that “fills in the gaps” in measured quantum properties is a continuity that is conceived or imagined; it cannot, with any possible meaning, be claimed (or denied) to possess possible reality beyond this. To claim (or deny) otherwise, as we have seen, is to fall victim to projection.

### 27.13 Quantum theory as a model of objectivity

In the course of this study, several distinguishable meanings of the general concept of objectivity have been discussed.

First, there is objectivity in the sense employed, for example, by Gaston Isaye ({6.7}): the objectivity of the fundamental and unavoidable commitments of rational discourse. This is the *objectivity of rationality*.

Second, there is objectivity as it is expressed by results that are not subject to controversy and which can be replicated by competent colleagues in a given discipline, provided they accept the framework which the establishment of those results presupposes ({10, 17.3}). This variety of *framework-relative objectivity* takes two forms: One is found in connection with formal systems in mathematics and logic, i.e., the *framework-relative objectivity of formal systems*: Given the premises or axioms of an individual formal system, logically necessary deductions from them are objectively valid in this sense.

A second form of framework-relative objectivity is found in connection with the natural sciences and is closely related to objectivity that concerns formal systems. It comprises a distinguishable third variety of objectivity. In the preceding chapter we encountered this variety of objectivity as it is defined by the invariance of physical laws with respect to a restricted set of observational frames of reference, as, e.g., in special relativity, or defined by the invariance of such laws freed from such restriction, as, e.g., in general relativity. This is the objectivity of principles of invariance which is sought by any natural, not purely formal, science, i.e., the *framework-relative objectivity of invariant natural principles*.

Fourth, there is objectivity defined as the correlation of perspectively given information ({21.1}). This is objectivity as it applies, for example, to the perceived compatibility of individual spatial frames of reference. Such perceived compatibility permits the correlation of multiple compatible perspectives so that reference is possible to objects considered to be the same from the standpoint of those different individual perspectives. This is *objectivity through the correlation of perspectives* or *correlation of reference frames*. It is the variety of objectivity that is involved in the concept of “external perceptual space”—the imagined, conceptualized, or constructed more



comprehensive reference frame in terms of which individual perspective spaces may be considered to form parts.

Fifth, there is objectivity in the sense in which we have encountered it in quantum theory, which at this point merits explicit analysis. This is objectivity with a complex meaning that combines, as we shall see, the foregoing four concepts of objectivity, supplemented by the additional concept of objectivity contributed by quantum theory, the *objectivity of complementarity*. This fifth variety of objectivity concerns the correlation of information given in what has earlier been termed ‘dimensionally incompatible reference frames’ ({10.5, 27.8}). This is the correlation of information pertaining to conjugate quantum properties obtained from the standpoint of mutually exclusive experimental arrangements.

In addition to these five important, familiar, and meaningful varieties of objectivity, there is a sixth variety of putatively meaningful objectivity which, as we’ve seen previously, is metalogically self-undermining. It is the purportedly meaningful notion of “objectivity” that we have met numerous times before, expressing a concern and desire for “reality that is independent of any framework of reference.” This is the notion of “objectivity” to which, as we have seen, Einstein was tenaciously and uncompromisingly committed ({21.1.2}). As we’ve also seen, it was his belief in such objectivity that put him at odds with quantum theory, leading him to refuse to accept indeterminacy and complementarity as “complete” descriptions of quantum phenomena. We have rejected this *projective notion of objectivity* in the course of earlier chapters and it will not concern us further in this chapter.

Finally, there is a separate, seventh distinguishable metatheoretical variety of objectivity, which is the goal of this study. This is a concept of objectivity that involves the universality and invariance of principles formulated from the standpoint of a self-referential, vertical, non-ordinal theory capable of studying all theories ({4}). I have called such a theory developed on the level of maximum theoretical generality the ‘metalogic of reference’; one of its central goals is to make possible reflexive proofs which demonstrate that certain claims cannot be denied without undermining their own possibility of reference and meaning. It comprises a “metalogic” in the sense that it formulates general principles which govern possible reference in any subordinate theory, and which also govern possible reference from its own standpoint. Principles and proofs that meet these requirements may be said to possess *metalogical objectivity*.

Of these seven distinguishable varieties of objectivity, the first five, the varieties likely to be familiar to all readers, have, as we shall see, become interrelated, indeed interwoven, in the seldom recognized and seldom fully

appreciated unifying concept of objectivity developed by quantum theory.

During the course of this chapter's discussion of quantum theory, I have brought together a number of claims that have been made, often uncritically, both by the original Copenhagen physicists and by some of their later interpreters. With the objectives of a "negative science" in view ({12-15, 23-24}), it has been my purpose to identify and to subject to de-projective analysis a group of these claims, each of which is metalogically self-undermining. A relatively small number of quantum theorists limit themselves to endorsing only very few of these claims, while most of these claims are accepted without hesitation by the majority of quantum physicists and by the majority of philosophers of physics. The following table summarizes many of these projective claims, which appear in the left-hand column, paired with their de-projected revisions, in the right-hand column:

<b>Putatively meaningful but projective quantum-theoretical claims</b>	<b>De-projected revisionary replacements</b>
Measurements bring about perturbing interactions between observers, their laboratory instrumentation, and quantum phenomena.	Measurements provide information about quantum phenomena: Their properties and our claims about them are limited to this information; the notion of "perturbing interactions" is not only without meaning, but is without possible meaning.
Properties may be ascribed to isolated quantum phenomena: For example, an electron "in reality" possesses a well-defined position or momentum independently of the precise measurement of the other.	Quantum phenomena are framework-relative: The properties of a quantum phenomenon cannot be separated from its functional relation to the dynamic system consisting of observer-measurement apparatus-quantum-object; no physical meaning can be associated with the view in the left-hand column.

<p>Information obtained through quantum measurements is essentially “incomplete.”</p>	<p>Information obtained by all the possible means afforded by the referential framework consisting of the theoretical system of quantum mechanics in conjunction with quantum laboratory measurements is complete information.</p>
<p>Quantum phenomena possess an autonomous reality, independently of the context established by individual observations and the conceptual structure of quantum theory itself.</p>	<p>The reality of quantum phenomena <i>consists in</i> their framework-relative functional interdependence in the dynamic system of observer–measurement apparatus–quantum-object.</p>
<p>Hidden variable theories are capable of providing an adequate account of quantum theory.</p>	<p>Hidden variable theories are referentially self-undermining because the referential preconditions they must meet cannot—in principle—be satisfied within the framework of the Copenhagen results.</p>
<p>What quantum theory observes is not nature itself, but nature as it is revealed by our interactions with it.</p>	<p>What is observed <i>is</i> nature; beyond this, the notion of “nature-itself” is without possible meaning.</p>
<p>Indeterminacy or uncertainty, randomness, and use of probability functions in quantum theory reflect deficient knowledge.</p>	<p>On a quantum level, nature is essentially indeterminate; randomness of many quantum events is characteristic of quantum reality; these facts about nature do not reflect deficiency of knowledge.</p>

**Table 27.1 Projective quantum-theoretical claims and their de-projected replacements**

Grouping together the de-projected results in the right column provides a summary description of the model of objectivity offered by quantum theory—that is, as quantum-theoretical objectivity is understood through the lens of the metalogic of reference. The resulting concept of objectivity accepts and affirms all five of the familiar forms of objectivity distinguished earlier:

- a. the objectivity of rationality
- b. the framework-relative objectivity of formal systems
- c. the framework-relative objectivity of invariant natural principles
- d. the objectivity through the correlation of perspectives
- e. the objectivity of complementarity

It is understandable that objectivity from the standpoint of quantum theory integrates all five varieties of objectivity: As we've seen, measurements of quantum phenomena presuppose a framework-relative context of functional interdependency consisting of the dynamic system of observer–measuring apparatus–quantum phenomena. This context is at once formal with empirical content (b. and c. above): The ability to identify invariant natural principles that describe the above functional interdependency on a quantum level in terms of acquired measurement information presupposes the formal conceptual framework provided by the mathematical formalism of quantum theory. Measurements of quantum phenomena presuppose that the distinct perspectives supplied by different laboratory set-ups can be correlated so that it is possible to judge that the same phenomena are in view (d.). Furthermore, it must be possible to judge that measurements of conjugate properties, undertaken by means of incompatible experimental arrangements, relate to what is judged to comprise the same quantum phenomenon, and this, in turn, presupposes the complementarity of the measurements (e.). Finally, the scientifically rigorous understanding of the foregoing measurements which have the objective of identifying invariant natural principles itself presupposes that the description of invariant quantum principles accords with standards of rationality (a.). We see that in each interrelated link of this chain, distinguishable forms of objectivity are presupposed.

The content of the previous paragraphs is descriptive of the standards of objectivity that are accepted by the still-dominant Copenhagen results. The objectivity of rationality,<sup>333</sup> (a.), is itself strongly compelling<sup>334</sup> since its rejec-

---

<sup>333</sup> On rationality and rules of admissibility, see {6 and 17}.

<sup>334</sup> For the meaning of 'strongly compelling', see {11.4}.

tion is metalogically self-undermining as well as pragmatically self-refuting. Objectivity in sense (b.) is frequently strongly compelling, while (c.), due to its empirical, contingent content, is compelling but not strongly compelling: Objectivity in senses (b.) or (c.) is expressed by results that are not subject to controversy and which can be replicated by competent colleagues in a given discipline, provided they accept the mathematical and physical framework which the establishment of such results presupposes. Quantum theory has contributed a revolutionary addition to standards of objectivity: As a result of its theoretical formulation and the experimental results obtained, the following have proved to be inescapably necessary to modern quantum theory: the correlation of perspectives, that is, of information acquired by means of different experimental arrangements, (d.), and the complementarity of information acquired from mutually exclusive experimental arrangements, (e.).

In the above interwoven sense, the “objective” understanding—and here I deliberately avoid the misleading term ‘interpretation’—which de-projected quantum theory provides both avoids and rejects *as essentially without meaning* the notions according to which:

- ◆ there exist, or do not exist, “measurement-based perturbations” resulting from “interactions” between observers, their instrumentation, and quantum phenomena
- ◆ properties can be predicated of “isolated” quantum phenomena
- ◆ quantum measurements are essentially “incomplete”
- ◆ quantum phenomena have “independent reality”
- ◆ hidden variables can be called upon to supply the “incomplete” information and restore “autonomous reality”
- ◆ quantum theory does not observe nature “as it really is”
- ◆ recourse to quantum probability functions is a sign of “imperfect” knowledge.

The objective concept of quantum reality that comes to light when these projective notions are eliminated, in keeping with the objectives of a “negative science,” provides a new and epistemologically sound understanding of quantum phenomena and their intrinsically defined basis: Quantum phenomena and that basis comprise a dynamic system which we have characterized as an interconnected totality consisting of observer–observer’s instrumentation–quantum object. It is important at this point that we recognize the usually overlooked fact that this dynamic system, however, includes somewhat more than these interrelated aspects:

We need to recall that when two frameworks of reference are dimensionally incompatible (for quantum theory, these are mutually exclusive laboratory set-ups), the two frameworks may be considered *complementary* when the set of objects identifiable in one frame is judged to comprise the same set of objects identifiable in the other. For this judgment to be made, as noted in {10.5}, a third framework, a meta-framework capable of correlating the two subordinate reference frames, must be presupposed. The metalogical need for such a third, more comprehensive frame of reference is seldom spelled out by quantum physicists, and yet it is the reference frame in terms of which the theory of quantum events itself must be articulated. The indispensable role of that meta-frame has to be recognized and incorporated in the dynamic system we have just characterized.

From the standpoint of that meta-framework, quantum reality is perspectivally given in a wide range of laboratory observational reference frames; quantum reality *is* the functional interdependency which the expanded dynamic system of reference comprises: that is, observer–observer’s instrumentation–quantum object–correlating meta-framework. *That total referential system defines quantum reality*, and it determines specific metalogical horizons which, if attempts are made to breach them, result in the projections summarized on the left-hand side of the preceding table. Reference to quantum objects involves a correlation of such perspectives, a correlation or *coordination* among perspectives which we discussed earlier in this study in connection with the concept that physical objects comprise “reference structures” which function in essentially coordinative terms.<sup>335</sup> The concept of objectivity provided by quantum theory provides a meta-level understanding of quantum reality as fundamentally and unavoidably coordinative in this sense. As we shall see in the next chapter, the quantum-theoretical integrative model of objectivity will be extended, on the basis of the results that the metalogic of reference has reached, to apply equally to macrophysical reality.

The resulting de-projected understanding of quantum reality requires that one relinquish and repudiate an entire group of conventionally accepted but metalogically self-undermining notions. The objective concept of quantum reality accepts that quantum reality is essentially characterized by discontinuity, indeterminacy, and randomness, a reality that is functionally interdependent as defined by the dynamic system that permits reference to it. Quantum phenomena are intrinsically such that mutually incompatible laboratory arrangements will provide information of conjugate quantum properties that is complementary, and this fact should be no more unsettling than the

---

<sup>335</sup> Readers will recall Reichenbach’s early suggestion in this direction when he introduced the incompletely developed concept of “knowledge as coordination.” See {5.7} and {26.1}.

example considered earlier of a stick in a glass of water.<sup>336</sup> Although from a visual perspective the perceived stick is bent, and although from a tactile perspective it is sensed to be straight, in terms of the unifying meta-perspective supplied by the theory of refraction we recognize that one and the same stick is involved. It is the same with respect to light: although from the standpoint of one laboratory apparatus it is recognized to be a wave phenomenon, and although from the standpoint of a different and incompatible laboratory set-up it is observed to be a particle phenomenon, from the unifying meta-perspective supplied by the quantum principle of complementarity, we judge that one and the same physical phenomenon is involved.

There is a conceptual simplicity in such a de-projected understanding. It is minimalist in its ontological and epistemological claims, and does not seek to transgress the horizons of its possible meaningfulness. I am reminded of a comment made by physicist John A. Wheeler (1990, p. 13), although made in a different context: “[W]e will grasp the central idea of it all as so simple, so beautiful, so compelling that we will say to each other. ‘Oh, how could it have been otherwise! How could we all have been so blind so long!’”

#### 27.14 Quantum theory as a set of limitative results

The recognition that the interaction between the measuring tools and the physical systems under investigation constitutes an integral part of quantum phenomena has not only revealed *an unsuspected limitation* of the mechanical conception of nature, as characterized by attribution of separate properties to physical systems, but has forced us, in the ordering of experience, to pay proper attention to the conditions of observation....

– Niels Bohr (1958, p. 74, italics added)

The chief intent of the methodology of the metalogic of reference is to identify limitations that cannot be exceeded without undermining the possibility of meaning; I have called these ‘metalogical horizons’. Although among quantum physicists we occasionally find statements that certain views or variant interpretations of quantum theory are without meaning, in most cases these statements are not justified theoretically. We find a more satisfactory example of such claims in Heisenberg’s statement, quoted earlier in a foot-

---

<sup>336</sup> See {10.5}.

note, that the “uncertainty relation specifies the *limits* within which the particle picture can be applied. Any use of the words ‘position’ and ‘velocity’ with an accuracy exceeding that given by [the uncertainty principle] ... is just as *meaningless* as the use of words whose sense is not defined” (Heisenberg, 1930, p. 15, italics added). Heisenberg makes clear in this statement that attempts to attribute to conjugate properties an accuracy that seeks to go beyond the limitative results of the uncertainty principle are devoid of meaning: Such attempts are without meaning precisely because they are, in the theoretical context determined by the uncertainty principle, ruled out as in principle impossible (and not simply because quantum physics lacks more sensitive instrumentation).<sup>337</sup>

In this chapter, I have brought together a wide range of quantum-theoretical claims which from the standpoint of the metalogic of reference must be judged to be without meaning. In each case, the justification that these claims are devoid of meaning is that each attempts to breach the limitative horizons that define what we’ve come to understand as the dynamic referential system of quantum theory. This justification does not appeal only to a prior acceptance of the uncertainty principle, as in the example drawn from Heisenberg above, but concerns, in addition and at least as importantly, attempted violations of the metalogical horizons that limit the possibility of reference with respect to the quantum phenomena under consideration.

The limitative results of quantum theory are of particular relevance to the metalogic of reference precisely because quantum theory has focused attention on the preconditions of physical reference to quantum phenomena. This attention has been directed specifically to the role of measurement, but, as we have seen, that attention has unfortunately been misdirected as a consequence of the projective notions that measurements on the quantum level “perturb” the phenomena measured “as a result” of “interactions” between the distinguishable observer, the observer’s instrumentation, and the objects observed.

Once we have corrected our understanding of the limitative preconditions of reference to quantum phenomena, the framework-relative understanding of quantum reality we obtain conforms in many respects with the framework-relative understanding reached in previous chapters in connection with macrophysical reality. We shall look briefly at possible “bridges” between the

---

<sup>337</sup> The same fundamental acknowledgment of meaninglessness has been made by Griffiths in expressing what he has called the ‘consistent histories’ approach to quantum theory: He refers to mutually exclusive experimental set-ups that can measure conjugate quantum properties as “inconsistent families of histories,” and about these families he claims: “Within the consistent history approach to quantum theory such families are ‘*meaningless*’ in the sense that there is no way to assign them probabilities within the context of a stochastic time development governed by the laws of quantum dynamics” (Griffiths, 2002, p. 142, italics added).



macro- and microphysical domains later in the next chapter. The explicitly framework-relative quantum-theoretical understanding of microphysical phenomena, coupled with the integrative concept of objectivity contributed by quantum theory, identify significant and important theoretically fundamental affinities that exist between the metalogic of reference and quantum theory. Both approaches are essentially limitative by delineating horizons beyond which meaningful claims, in principle, cannot be made.

## Epistemological Lessons Learned from and Applicable to Relativity Physics and Quantum Theory

In studying the work of the group of physicists who were the founding fathers of relativity theory and quantum theory, one cannot help but be struck by the surprising extent of their knowledge of philosophy. Many, including Einstein, Planck, Bohr, Heisenberg, von Weizsäcker, and others, studied philosophy while in school, and then throughout their professional careers as physicists maintained an active interest in reading philosophical works and in self-consciously relating their ideas and approaches to philosophical issues. The greatest portion of their philosophical concerns specifically involved epistemology, and many of them looked to epistemology for clarification, if not the resolution, of fundamental problems they wished to solve.

When Newton published his 1687 *Mathematical Principles of Natural Philosophy and the System of the World*, physics had not yet separated from philosophy. More than two centuries later, when Einstein, Bohr, and Heisenberg made their major contributions, philosophy, and explicitly epistemology, still continued to exercise significant influence on physics. In {26.3} I described as fundamentally epistemological Einstein's approach in formulating invariant physical principles. He described the objectives of physics in unmistakably philosophically focused terms. His conceptual path to the special and general theories of relativity was, as Lorentz and others remarked, explicitly epistemological, asserting a primary concern for "first principles" and logically compelling deductions that can be made from them. We recall that Einstein claimed: "Science without epistemology is—in so far as it is thinkable at all—primitive and muddled" (Einstein, 1959/1949a, p. 684), while Lorentz, referring to special relativity, remarked "[t]he acceptance of these concepts belongs mainly to epistemology (Lorentz, 1913, p. 23).

Very similarly, Bohr characterized his approach to quantum theory explicitly in terms of conceptual analysis with an epistemological focus: "[E]mphasis is laid especially on the presuppositions for unambiguous use of the concepts employed in the account of experience.... We are here faced with

an epistemological problem quite new in natural philosophy...” (Bohr, 1958, pp. 2, 25). In describing his work, Bohr repeatedly made mention of the “epistemological lesson” that influenced the development of quantum theory (Bohr, 1958, pp. 74, 88). As will be made clear, a number of such epistemological lessons have been involved, so in this chapter I use ‘lessons’ in the plural.

Throughout the history of philosophy and especially the history of philosophy of science, philosophers have paid careful attention to major developments in physics and have subjected these to philosophical analysis, albeit sometimes only after long delays. Although today we do not find many theoretical physicists with the strong philosophical backgrounds evident among major physicists a century ago, physics and epistemology have not entirely gone their own ways. There continues to be a certain degree of exchange of ideas between the two disciplines, and some cross-fertilization of ideas. Unfortunately, most especially in the United States in contrast to Europe, the tendency of higher education has reduced the amount of exposure to philosophy that physicists now receive during their university training, while the degree of increasingly specialized and technical scientific knowledge required of philosophers in order to engage in fruitful work in philosophy of science has inevitably placed an unfortunate and increasing distance between the two disciplines.

What I’ve described in the preceding paragraphs has related largely to *exchanges* between physics and philosophy; occasionally there have been *influences* of the first upon the second; but seldom during the past hundred years and more have we found work in philosophy that exerts a significant transformative effect upon the course of physics.

There are two major ways in which philosophy of science is able to contribute in a conceptually major sense to work in physics. In describing these, from this point on I shall limit what I have to say with epistemology expressly in view. The most familiar way in which epistemology has been able to contribute to the development of physics is by offering meaningful, systematic, reflective clarification of important and fundamental ideas that have been developed by physicists, which are relied upon in the course of their work, and which are employed in the communication of what they claim to know on the basis of that work.

A second and quite different way in which epistemology can contribute to the course of physics is by means of the results reached by the negative science of the metalogic of reference. Such results complement the positive science of physics by identifying the horizons of meaning and possibility which the frameworks of physical theories determine. Where the domain of positive

science is defined by empirical theories that have been developed, by experimental results that have been reached, and by claims that can empirically be confirmed on their basis, the corresponding portion of the domain of negative science that concerns physical science is defined by its self-validating metatheory and the metaclaims derived from it concerning the horizons of positive science which negative science seeks to identify. The results of the positive science of physics attempt to describe invariant principles, the laws of nature, which govern empirical reality, while the corresponding results of negative science identify the parameters of constraint which define the system of reference that constitutes each particular positive theory, parameters of constraint which define therefore what is possible from each standpoint, and hence what is possibly meaningful from that standpoint. In this work, negative science plays the role, not of Locke's handmaiden to positive science, but the role of a *complementary, logically symmetric, and necessary equal partner*.

Physicists, no less and no more than ordinary mortals, are prone to conceptual excesses and distortions of both their theoretical and empirical understanding and the communication of that understanding, excesses and distortions that come about when they inadvertently or willfully seek to transgress beyond the horizons of possible meaning which their own concepts and theories metalogically entail. The "critique of impure reason" undertaken in this study represents an indispensable step to remedy such transgressions. As we've seen, its general aim is to detect and reject concepts and beliefs based on them that are metalogically self-undermining, and then to replace these concepts and beliefs with corresponding revised, non-projective concepts and meaningful claims. This is self-evidently a task of conceptual criticism, evaluation, and revision. It is a task, as the early chapters of this book have pointed out, that often runs against the familiar grain and can present a challenge to those not willing to relinquish their preferred beliefs. Despite the challenges posed by human psychology, *positive science without its negative, critical counterpart is only half-evolved*.

In the course of the last two chapters concerned with relativity physics and quantum theory, we have mentioned, but have not made explicit, two principal distinguishable routes to the recognition and confirmation of framework relativity. On the one hand, there is the route taken by the first 25 chapters of this study, the majority of which have not focused on questions posed by physics; on the other hand, there is the route taken by physics in its development of relativity theory and of quantum theory. Each route constitutes an approach that has confirmed framework relativity: The first route, defined by the metalogic of reference, has been developed wholly on an abstract level, without necessary reference to empirical content, with the intent of describing

invariant metalogical principles that govern the possibility of meaningful reference. The de-projective method and the wide variety of its applications involve, as we've seen, purely theoretical, conceptual analysis.

The second route, defined by the special and general theories of relativity and by quantum theory, has an unmistakably strong and thoroughly conceptual component, but one which acquires physical meaning through applications to the world of physical phenomena, and through those applications the three physical theories are judged successful to the extent that they provide satisfactory descriptions of physically invariant principles.

These two routes, different though they are, interrelate on a purely theoretical level which we might call '*mutual confirmation*'. It is in the relation of mutual confirmation that we find the most significant epistemological lessons from which each can learn from the other.

### 28.1 The relation of mutual confirmation

If we consider ways in which the reflexive metatheory developed in this book "confirms" the framework-relative results reached by both relativity physics and quantum theory, we need to recognize that such a confirmation relation is effectively one-way: The methodological results of the metalogic of reference are formulated, as we've seen, on a level of maximum theoretical generality. They are abstract, generalized results that enable us to identify parameters of constraint upon possible reference from the standpoint of any theory, including the metalogic of reference itself. The concept of metalogical horizons has been developed in this metatheoretical context. As we have seen, the concept of horizon is an essentially modal concept because it pertains to the individual limits of possible reference associated with any given theory. To the extent that, for example, special and general relativity confirm, bear out, or ratify certain of the framework-relative principles reached by the metalogic of reference, then to that extent special and general relativity, from the standpoint of the negative science of the metalogic of reference, *comply* with the self-validating principles to which the methodology of the metalogic of reference has led.

It should be evident that solely because a theory complies with the prescriptive horizon-limitations identified by negative science, it does not follow that such a theory is in any sense, beyond that fact, validated or shown to be an adequate theoretical description of the domain of objects for which it has been developed. A theory's compliance with its own metalogical horizon is a necessary condition of its possible reference and possible meaning, but that compliance is not of course sufficient to substantiate the theory itself. Ex-

pressed differently, an individual theory, e.g., special relativity, may, in important respects, “confirm” or bear out the metalogical principles of the metalogic of reference insofar as special relativity in many instances complies with those principles, but this relation of confirmation extends in only one direction. The metalogic of reference is not itself in any significant sense “confirmed” by the fact that special relativity, in certain ways, conforms to its principles, although one may feel heartened that relativity physics has made a revolutionary advance in large measure due to Einstein’s observance of principles of framework relativity. Nonetheless, in the latter sense we may say that, to the extent that special relativity does bear out certain of the principles established by the metalogic of reference, then in those respects the results reached by special relativity are “confirmed” or “validated” by the metalogic of reference.

In the vocabulary of logic, conformity with the principles of the metalogic of reference is a necessary but not sufficient condition for the validity or acceptability of an individual theory. The capacity of an individual theory to refer to an intended range of objects of reference does depend, necessarily, upon such compliance. When an allegedly meaningful individual theory or presumably meaningful results to which that theory leads fail in such compliance, to that extent the theory or its results attempt to trespass beyond that theory’s horizon of possibility and meaning, a fact which then provides strongly compelling justification that it is necessary to revise or reject the theory itself or its projective content.

We may see in retrospect that the two preceding chapters concerned with relativity physics and quantum theory have focused on the relation of mutual confirmation that centers attention on the physical theories themselves, identifying how those theories bear out—but also sometimes fail to bear out—principles of framework relativity to which the metalogic of reference has led. As remarked in the Preamble preceding {26}, no other natural scientific theories exemplify the degree of disciplined analysis of physical framework relativity as do relativity theory and quantum theory. To the extent that I have given space to a discussion of these physical theories in this work, we have observed, in a variety of ways, how relativity theory and quantum theory exemplify comparatively consistent rigorous adherence to the theoretical constraints of framework relativity. As we shall summarize in subsequent sections of this chapter, both “confirm” in the sense employed here certain of the main results to which the metalogic of reference leads. In these ways, relativity physics and quantum theory successfully, in the specific ways I identify, pass through the error-detecting “sieve” ({12}) of the negative science of the metalogic of reference: Both their principal results and the theoretical meth-

odology employed to reach those results comply in great measure with the constraints which the metalogic of reference seeks to enforce. As anticipated in the Preamble, both relativity theory and quantum theory have attempted to go to the heart of framework-relative analysis within the context of theoretical physics; both approaches are distinguished by their sometimes only implicit, but often explicit, analysis of the preconditions of mathematical and physical reference to the kinds of objects—macroscopic in one case, microscopic in the other—which comprise their objects of study. Framework relativity is not only accepted and explicitly endorsed by both approaches, but principles of framework relativity are firmly woven into the conceptual and mathematical formulation of both theories, although, as we have found, this is true to a lesser and less satisfying extent in much of their ordinary language expression and interpretation.

From a more critical point of view, we have seen that neither relativity theory nor quantum theory is entirely free of self-undermining projective claims. In the last two chapters, an effort has been made to show how it is possible to de-project a number of those problematic assertions. The fact that relativity physics and, to a greater extent, quantum theory sometimes fall short of full consistency with metalogical principles of coherent reference brings to light areas of both theories that can profit from de-projective analysis in order to insure that both theories respect their own metalogical horizons. It is here that physics can learn most from the negative science of the metalogic of reference.

In the next sections, we consider ways in which relativity theory and quantum theory have profited, and can still profit, from epistemology and, in particular, from the approach of the metalogic of reference.

### **28.2 The main epistemological lessons learned from and applicable to relativity physics**

The main objective of a negative scientific analysis of an individual theory is the identification and correction of that theory's concepts and claims that undermine their own possibility of meaning. When correction is not possible, then such projective concepts and claims must be rejected. As we've seen, to accomplish such a negative scientific analysis, the preconditions of reference of the theory in question must be made explicit. It was in just such an analysis of the preconditions of physical reference that Einstein excelled.

The main epistemological lessons from which his work profited can be divided into two categories: *renunciation* of traditional notions of classical physics which had largely been accepted uncritically by earlier physicists, and

*affirmation* of principles of framework relativity. In {26}, Part I of Table 26.1 summarizes many of the claims of classical physics which Einstein's painstaking and referentially focused analysis forced him to renounce. These renounced claims included: The interval of time separating two events is independent of the state of movement of any frame of reference—i.e., the duration of an event is an absolute quantity; the simultaneity of two events determined from the standpoint of an inertial system implies the simultaneity of the events in relation to every inertial system; the spatial distance between two points on a rigid body is independent of the state of movement of the reference frame of measurement—i.e., the length of a rod is an absolute quantity; there is a privileged system of reference; movement is an absolute quantity—i.e., a moving object defines an absolute, single, unique trajectory; absolute space and time exist; empty space exists; the acceleration of an object and its response to gravitational field forces are fundamentally distinct; in gravitational fields rigid bodies with Euclidean properties exist; etc. In each of these instances, it was Einstein's disciplined analysis of what I have called 'preconditions of physical reference' that led to his renunciation of this large and influential group of classical notions. And in each of these instances, as we have seen, it is also the case that a metalogical analysis of the same classical claims leads to their rejection. The rationale for their rejection differs as a consequence of the dissimilar methodologies of relativity theory compared with that of the metalogic of reference, and that rationale differs also as a result of distinct levels of abstraction and theoretical generality.

Part II of the same table summarizes the methodological principles Einstein followed in developing the special and the general theories; two of these, (j) and (k) in the table, crucial to the development of the special theory, explicitly concern framework relativity. The two principles affirm that the representation and formulation of physical laws must accept the framework relativity of physical phenomena, and that the physical identity of phenomena—understood in terms of their measurable properties—is essentially a function of the reference frame employed to identify them. These principles, too, as we have seen, have their parallel correlates in the metalogic of reference.

Two other methodological principles listed in Part II of the table also concern framework relativity, but framework relativity understood in an intrinsic, non-relational form which can be correlated with its associated form in the metalogic of reference. Those two principles, (n) and (o) in the table, were these: From the extrinsic nature of pre-relativity frameworks of reference, Einstein made a shift to intrinsic mathematical means of identifying physical phenomena; to make such a shift possible, he applied the mathematical tools



of tensor analysis which made it possible to formulate, in wholly intrinsic terms, functional interrelationships among physical phenomena. Here, as well, there is a parallel correlation with the metalogic of reference: As we've seen, in the metalogic of reference this takes the form of the non-relational, intrinsic embedment of the constitutive structure of a given framework of reference in the very identifiability and identities of objects to which that framework is capable of referring.

The preceding paragraphs comprise a greatly abbreviated outline of ways in which relativity physics has incorporated negative results by renouncing notions and claims of classical physics, while it has at the same time affirmed positive principles of framework relativity that follow from an epistemological analysis of referring systems. As we consider Einstein's work in retrospect, these negative and positive insights identify fundamental ways in which relativity theory may be thought to have learned from epistemology. Einstein was, it may be surmised, implicitly aware of some of these lessons from epistemology, but the influence of other lessons we are able to recognize only in a perspective of interpretation as we consider his achievement in retrospect.

However, like many important contributions, Einstein's special and general theories of relativity, as well as his subsequent research in physics, suffered from what we have seen was his self-undermining beliefs in the purportedly meaningful notions that quantum phenomena "exist independently of their observability," that between successive observations of the state of a quantum system, "continuously existing physical reality must nonetheless 'be there' "—despite quantum randomness, discontinuity, and the essential need to accept that the predictability of quantum events is governed by probability distributions. Einstein was, furthermore, unable to renounce classical causality on the quantum level; he was unable to accept, as a result of the inescapable framework relativity of dynamic quantum systems, that it was also necessary to let go of the classical notion of causality.

In each of these cases, as we saw in {26}, analyses provided by the metalogic of reference make clear that all of these beliefs must be relinquished because all are metalogically projective, undermining the possibility of their meaning. In all of these cases, relativity physics is still able to profit from the lessons of epistemology, and specifically from the negative science of the metalogic of reference.

### **28.3 The main epistemological lessons learned from and applicable to quantum theory**

Notwithstanding all differences between the physical problems which have given rise to the development of relativity

theory and quantum theory, respectively, a comparison of purely logical aspects of relativistic and complementary argumentations reveals striking similarities as regards the *renunciation* of the absolute significance of conventional physical attributes of objects.

– Niels Bohr (1959/1949, p. 239, italics added)

[T]he resolution of the paradoxes of atomic physics can be accomplished only by further *renunciation* of old and cherished ideas.

– Werner Heisenberg (1930, p. 239, italics added)

Also due in large measure to the clarity offered by retrospective understanding, we are able to recognize very much the same benefits of epistemological lessons at work in the thought of the pioneering contributors to quantum theory as we did in connection with relativity physics. As in the case of Einstein's special and general theories, the development of quantum theory involved both the renunciation of notions and claims of classical physics, and the affirmation of principles of framework relativity.

To separate these clearly, first an abridged statement will be given of a group of the principal claims embodied in the Copenhagen results as they most commonly continue to be expressed today. This is followed by an equally abbreviated review of the main renunciations by the Copenhagen physicists of classical claims. The foremost metalogically projective claims which were associated in the last chapter with the Copenhagen results are then summarized, followed by a description that brings together the foregoing claims in a de-projectively revised and corrected re-statement of the original Copenhagen results. That re-statement embodies the epistemological lessons which, to varying degrees, have been learned—and can yet be learned—by quantum theorists.

In the preceding chapter, we saw that the Copenhagen results involve the following principal claims: There is an unavoidable interaction between the act of measurement of quantum phenomena and the objects measured; the act of measurement changes the state of quantum phenomenon measured. This is the notion of measurement-based perturbation. There is, furthermore, an absolute lower limit of accuracy to which conjugate quantum properties can simultaneously be known. This is Heisenberg's principle of indeterminacy or uncertainty. This limitative result is incorporated in the recognition that, rela-

tive to mutually exclusive experimental set-ups, incontrovertible observations are made of mutually exclusive quantum phenomena, e.g., the wave nature of light and the particle nature of light. The acceptance that both mutually exclusive quantum properties reflect a fundamental fact of quantum reality is expressed by the principle of complementarity. The nature of quantum phenomena is in many respects essentially indeterminate, discontinuous, and random, and as a consequence can only effectively be described in terms of probability functions. The foregoing Copenhagen results provide a complete description of quantum reality; there is no meaningful sense in which they can be judged to be incomplete or to reflect deficiencies in physical knowledge that can, in principle, be overcome.

The acquiescence of physicists to the above results was, as we've seen earlier, sometimes difficult, and sometimes out of the question as it was for Einstein. The full acceptance of the Copenhagen results required, as it did in the case of special and general relativity, the renunciation of a group of cherished classical *ideals*. These ideals were: Microphysical objects, like their macrophysical counterparts, are believed to exist independently of observation, with an autonomy unaffected by whether or not they are observed, and, like their macrophysical counterparts, are believed to continue to exist during discontinuous periods of observation. This ontological claim expresses the classical notion of physical objectivity. Furthermore, it was the classical ideal that acts of quantum measurement do not necessarily interfere with precise measurements of the objects measured; measurements of quantum phenomena need not disturb or render indeterminate the properties of the measured phenomena. There is no necessary lower limit of precision to which quantum properties may simultaneously be measured. Mutually exclusive experimental arrangements designed to measure conjugate quantum properties will yield different results, but nevertheless quantum phenomena possess an underlying, unitary, independent reality that should ideally be characterizable in terms of non-contradictory properties that possess a reality autonomous of any and all acts of measurement. Indeterminacy, discontinuity, and randomness on the quantum level are indications of the unsatisfactory, incomplete state of current quantum knowledge, which, if and when appropriate advances are made, will eventually result in deterministic predictability of continuous and non-random quantum phenomena.

The resulting conflict over these hard-to-relinquish classical ideals, epitomized by the impassioned controversies that have occurred in the aftermath of the 1935 Einstein, Podolsky, Rosen paper, in what will soon be the passage of a century, has still not been resolved in a manner with which all quantum physicists agree. It is here that the negative science of the metalogic of refer-

ence can provide assistance in the form of strongly convincing problem-resolution. The application of the methodology developed in this study to the cherished classical claims summarized in the preceding paragraph has shown each claim to be self-undermining on the metalogical level—that is, in terms of an analysis of the preconditions of reference that would need to be granted in order for each of these classical claims, in principle, to possess possible meaning within the framework of established quantum theory. In analyzing each claim, we have found that each is projectively self-undermining; each seeks to assert meaningfully that which, given the quantum-theoretical framework within these claims are formulated, cannot, in principle, meaningfully be asserted.

Each of the following putatively meaningful notions and claims has been found to be projectively meaningless: The ontology of quantum phenomena is a realist ontology—that is, the objectivity of quantum claims requires that quantum phenomena possess an “existence autonomous of observation”; active “agency” is involved in quantum measurements; “perturbation” occurs “as a result” of such measurements; the observer, the observer’s instrumentation, and the observed quantum phenomena can all be “separably distinguished” from one another; the Copenhagen results are “incomplete,” reflecting the present state of “imperfect knowledge” on the part of quantum physicists; quantum reality is “continuous, causally deterministic, with well-defined properties that can be measured to an arbitrary degree of precision”; etc. —The recognition that each of the foregoing claims is metalogically self-undermining and therefore devoid of meaning clearly is an epistemological conclusion. That conclusion is not of course a result derivable from physical theory, nor is it implied by experimental data. It is neither of these, but rather a result that follows from an analysis of the most basic preconditions of reference and meaning with which any coherent, rational theoretical account must, in principle, comply, whether it is a theory of natural phenomena or of an altogether different domain of objects.

The following is a concise description that brings together the foregoing claims in a de-projectively revised and corrected re-statement of the original Copenhagen results: Quantum reality is intrinsically correlational in nature: This means that what is *real* on the quantum level *consists* in correlations among a plurality of possible observational perspectives. Having rejected as without possible meaning the notion that quantum phenomena possess a form of existence that is autonomous of observation, quantum reality is recognized to be functionally interdependent in relation to the combined formal and observational framework in terms of which they can in principle be observed. The dynamic system formed of observer–observer’s instrumentation–quantum

events—correlating meta-framework defines and *is* the contextual reality studied by quantum physics. Every quantum event is an event given in the referential field established by such a dynamic system. Measurements of conjugate quantum properties are complete and exhaustive of possible information. Quantum reality is essentially characterized by complementarity, indeterminacy, and the fact that quantum level descriptions take the form of probability functions. Attempts to characterize the physical state of a quantum system between observations are devoid of meaning, although “intervening” potential states can be given abstract expression mathematically. Quantum theory, understood de-projectively, is above all a set of a *physically limitative results*. The framework established by those results is *objective* in a new sense that replaces the traditional notion of objectivity with an integrative concept of objectivity that combines the objectivity of rationality, the framework-relative objectivity of the mathematics of quantum theory, the framework-relative objectivity of invariant natural principles, the objectivity resulting from the correlation of observational perspectives, and the objectivity of the complementarity of quantum phenomena that exhibit conjugate properties under mutually exclusive laboratory conditions.

It is often said that it is difficult to make sense of the principal results of quantum theory and that the theory gives rise to paradoxical consequences. Yet, from the de-projective standpoint there are no “paradoxes” engendered by quantum theory. From the standpoint of the metalogic of reference the Copenhagen results make very good sense, are natural and acceptable consequences of framework relativity, and are non-paradoxical. When regarded from the standpoint of the metalogic of reference, it is hard to understand why those results should be thought objectionable, puzzling, or mysterious.

The still-ongoing disagreements among both quantum physicists as well as among philosophers of physics relating to the projective notions and claims we have listed, are disagreements capable of final resolution—provided that the participants in these disagreements are willing and able to profit from the lessons of negative science.

#### **28.4 The main epistemological lessons to be learned from relativity theory, quantum theory, and the metalogic of reference**

Epistemology has a tendency to lag behind the actual development of science.

– C. F. von Weizsäcker (1980/1971, p. 194)

It should by now be evident that epistemology has contributed in fundamentally significant ways to the development of both relativity physics and quantum theory. That this is the case was explicitly recognized, as we've seen, by Einstein, Bohr, and Heisenberg. This chapter and the two chapters preceding it have identified a substantial number of ways in which epistemological analysis served to guide their thought. These have underscored the influence of epistemology upon relativity theory and quantum theory, rather than the reverse, the influence of relativity theory and quantum theory upon epistemology. It is about this direction of influence that I would like to add a few words here.

There are, to be sure, lessons in epistemology that can be learned from the source of theoretical physics and to the benefit of epistemology, as well as to the benefit of general philosophy. As in the instances of both relativity theory and quantum theory, the epistemological lessons that can exert the greatest influence upon philosophy are of two kinds: one involves the renunciation of many traditional philosophical notions, notions which have largely been accepted uncritically, and the other involves the affirmation of principles of framework relativity. In both of these ways, epistemology, as well as philosophy generally considered, are able to profit.

As I appraise the discipline of philosophy, it has been clear, very much unlike incremental and progressive evolution in science and particularly in physics, that philosophy, including epistemology, has failed to find a common, shared, and rationally compelling methodological framework in terms of which contributions to the discipline can be constructively criticized and evaluated. There is a great and as yet largely unfelt need by philosophers to undergo the same kind of disciplined, rigorous purging of traditional, putatively meaningful notions and claims, a renunciation which has so profoundly contributed to the advancement of physics. But because of the absence of a unitary, agreed-upon methodology and universally accepted standards of validity, the renunciation of traditional notions and claims of philosophy faces stronger resistance than in physics.

The greatest number of chapters in this long book have had to do with concepts and claims that have, for many centuries, formed the center of attention of epistemologists and of philosophers of all stripes. The method of de-projection developed in this study has been applied in previous chapters to one after another of these concepts and claims, chosen because each has, in its own way, attempted to transgress horizons of possibility and meaning, and, as a consequence, undermines its own possible meaning. The negative science to which the metalogic of reference seeks to contribute provides a framework from which a comprehensive critique of impure reason can be undertaken.

Each of the concepts and claims shown in previous chapters to be projectively self-undermining has formed part and parcel of much of the traditional conceptual vocabulary of both epistemology and general philosophy. De-projectively corrected revisions of these concepts, when this is possible, have been given, and when not possible, their rejection has been urged.

In retrospect, there are clear lessons that epistemology as well as general philosophy can learn from the epistemologically insightful advances made by the physics of relativity and by quantum theory. To learn from these lessons could enable philosophy to make progressive steps in epistemological thinking to catch up with advances in the epistemology of theoretical physics. Nonetheless, such lessons are, one might say, “second-hand lessons” since they would be acquired from the successes realized by physics. However, there are, in addition, more general and direct lessons from which epistemology and philosophy can benefit, and these include the lessons provided by the metalogic of reference.

### **28.5 The wider applicability of these epistemological lessons**

We are ... dealing ... with an investigation of the conditions for the proper use of our conceptual means of expression. Such considerations not only aim at making us familiar with the novel situation in physical science, but might on account of the comparatively simple character of atomic problems be helpful in clarifying the conditions for objective description in wider fields.... [T]he straightforward solution of the unexpected paradoxes met with in the application of our simplest concepts to atomic phenomena might ... help us to clarify conceptual difficulties in other domains of experience.

– Niels Bohr (1958, pp. 2, 20)

In this third part of the book devoted to applications of the metalogic of reference, I have given more space to relativity physics and quantum theory than I have to other topics. I have done this for three reasons: One is that relativity physics and quantum theory have, in their own terms, a good deal to teach us that is directly relevant to the purposes and conclusions of the metalogic of reference.

The second reason is that both relativity theory and quantum theory bear out many of the results that the metalogic of reference has reached by means

of a wholly theoretical, non-physical method. This fact does, I would hope, help the skeptical reader to recognize that the results of the metalogic of reference are not a group of wholly abstract, far-removed, and inapplicable conceptions, but are rather conceptually fundamental and necessary principles without which it is possible to reach a coherent understanding of reality.

The third reason is that quantum theory deals with the microphysical domain, and relativity physics with the macrophysical universe: They are studies of both ends of the spectrum of physical phenomena. Many physicists have suggested that a unifying bridge ideally may be developed to connect the physics of the very small with the physics of the very large. Several designs for such a bridge have been contemplated by physicists: a unified field theory, the hope that inspired Einstein; relativistic quantum mechanics and its progeny, including string theory; extensions of quantum theory so as to include the macrophysical world;<sup>338</sup> as well as other proposals. These visions of a possible comprehensive and unitary theory would seek to describe invariant principles that govern the full range of physical phenomena. It has been and is a worthy ideal for physics.

However, there are, to be sure, many classes of phenomena that do not fall within the scope of physics. In contrast, the range of application of the negative science of the metalogic of reference extends to any rationally coherent attempt to identify and understand sets of objects of reference, of whatever kind they may be. The idea of a negative science defines a unifying bridge among possible theories by identifying the metalogical horizon of any given theory, making clear the limitative boundaries of its possible meaning.

We have seen how certain of the central results of special and general relativity and of quantum theory can be “confirmed”—in the singular sense described earlier—by means of the purely theoretical, non-physical method of de-projection. Yet the unifying results of this method extend beyond physics and pertain to any coherent rational framework in terms of which claims are made that are meaningful. Here, the reader is reminded that, for the metalogic of reference, the concept of meaning is open-ended. But it is not thereby unlimited or unrestricted, for referential consistency, in the sense defined by this study, remains the strongly compelling criterion, which cannot *not* be

---

<sup>338</sup> For example, Griffiths (2002, p. 368): “Quantum mechanics is clearly superior to classical mechanics for the description of microscopic phenomena, and in principle works equally well for macroscopic phenomena.”



accepted without becoming self-undermining on the level of possible meaning.



# **PART IV**

## **HORIZONS**

What is wanted ... is some new effort of logical imagination,  
some glimpse of a possibility never conceived before....

– Bertrand Russell (1972/1914, p. 245).

It is as if I were to try to explain the new ideas of any age  
to a person of the age that has gone before.

– Ford Madox Hueffer [Ford Madox Ford]  
and Joseph Conrad (1901, p. 10)



## 29

### Beyond Belief

*Virtus est vitium fugere et sapientia prima stultitia caruisse.*  
(It is the beginning of virtue and wisdom to flee from vice  
and free oneself from folly).

– Horace (*Epistle I, 1, 41hl f*)

[Y]ou ... talk about the truth, which turns out to be nothing  
but what you like to believe.

– George Bernard Shaw (1934, p 182)

#### 29.1 The negative science of the metalogic of reference

We may distinguish two fundamentally different but complementary approaches to scientific investigation, one familiar and the other not. One is exemplified by the natural sciences and mathematics, which I shall group together and refer to as ‘positive science’.<sup>339</sup> Positive science attempts to find invariant principles which govern natural phenomena in the case of the natural sciences, and which govern purely formal objects and structures in the case of mathematics. The less familiar approach to scientific investigation is exemplified by the metalogic of reference. It proposes a scientific study of the inevitable limitations that come with the adoption of any coherent frame of reference, whether it is the frame of reference of a theory or the frame of reference presupposed by a concept or claim. The scientific study of such limitations is not a familiar task associated with existing natural science or mathematics, but it is a task fundamental to a reflective understanding of concepts, claims, and their role in theories of positive science, and their role in philosophy and in other disciplines committed to rational coherence. This is a task

---

<sup>339</sup> There is no implied reference to Auguste Comte’s notion of “positive” philosophy or to positivism.

that is fundamental in order that we may know in what ways such concepts, claims, and theories are necessarily limited in what they can possibly refer to, what they can possibly mean, and—as a result of the foregoing—what can possibly be known by their means.

I use the term ‘science’ to designate an approach that is defined by a clearly formulated methodology and a set of rigorous standards of demonstration, a methodology and set of standards that can be communicated among multiple researchers and employed by them in ways that are able to be subjected to uniform evaluation and judgment through the application of that shared methodology and standards of demonstration.

A major objective of Kant’s *Critique of Pure Reason* was to provide an account of the conditions of *possible* knowledge, and by doing this to establish the boundaries of knowledge, beyond which, in Kant’s view, lies the domain of metaphysics. The present *Critique of Impure Reason*, in parallel but also in contrast, seeks to provide a *negative* account of those conditions which, if transgressed, result in *meaninglessness* and the *impossibility* of knowledge. It is a critique of “impure” reason because its purpose is to identify such transgressions of the limitations of reason, and by doing this to recognize a fundamental and widespread source of conceptual dysfunction that I have called ‘metalogical projection’. A critique that sets for itself the task of identifying and eliminating such “impurities” may suitably be called a ‘critique of impure reason’.

The metalogic of reference developed in this *Critique of Impure Reason* is characterized by a clearly formulated methodology and a set of rigorous standards of demonstration which can be communicated, shared, and evaluated by multiple researchers. It comprises what I have called a ‘*negative science*’. The conditions which become the focus of study of such a critique of impure reason are the referential preconditions that must be satisfied by any rationally coherent concept, claim, or theory. Its approach is “negative” in that it is concerned with what is referentially *forbidden*—on pain not only of meaninglessness, but of the *impossibility of meaning*. Its objective is to identify and to reject concepts and claims that attempt to transgress the limits of possibility and meaning, which in this study have been called ‘*metalogical horizons*’. The task of identifying and eliminating that which undermines its own possibility of reference is clearly a *negative* enterprise. A rigorous approach that accomplishes this end may appropriately be called ‘negative science’. Its methodology and its results are essentially *limitative*, in contrast to the results of the positive sciences whose goal is the description of invariant principles governing the behavior of their respective classes of empirical or purely formal objects of reference. In very general and over-simplified terms, we might

say that the concern of the *positive sciences* is the identification of principles that govern *what is the case or what is true*, whereas the *negative science* of the metalogic of reference seeks to identify principles that govern *what cannot be the case*, because to claim otherwise entails metalogically self-undermining reference that is meaningless. “Negative science” understood in this sense clearly sets a negative task.

I have chosen this phrase in part due to its use by Kant in a letter that he wrote in 1770 to Johann Heinrich Lambert; some brief comments about this letter were made earlier in {12}. Kant’s notion of negative science as he expressed it to Lambert is suggestive, but not of course fully developed within a single letter. As I will make clear, the negative science of the metalogic of reference should *not* be equated with Kant’s notion of negative science; it is important for readers to note that I do not use the term ‘negative science’ in order to carry out the purposes for which Kant appears to have intended it. I need to emphasize this—especially for Kant scholars. The reasons I’ve used this phrase is because it fits the tasks to which I put it, and because these tasks can be characterized as fundamentally in keeping with the objectives of transcendental argumentation as described in {9} and elsewhere in this study. It will, however, be useful to bear in mind what Kant had to say about “negative science” in his letter to Lambert. Let us place his comments in their historical context.

Johann Heinrich Lambert (1728–1777) was a Swiss-German mathematician, physicist, and philosopher. His contributions in mathematics included the first proof that  $\pi$  is irrational, early conjectures relating to non-Euclidean space, and a pioneering attempt to develop a calculus of logic. His main work in philosophy considerably influenced Kant; its title was *Neues Organon oder Gedanken über die Erforschung und Bezeichnung des Wahren und dessen Unterscheidung vom Irrthum und Schein* [*New Organon*,<sup>340</sup> or *Thoughts on the Study and Denotation of What Is True and Its Distinction from Error and Illusion*]. As the title indicates, the main objective of Lambert’s book was to establish the means to separate truth from error and illusion. It was likely this very concern in Lambert’s work to formulate ways to identify and avoid error that attracted Kant. Kant thought very highly of Lambert: “[I]n all sincerity, I hold you to be the greatest genius in Germany.”<sup>341</sup> He was sufficiently impressed and persuaded by Lambert’s approach that he planned to dedicate the *Critique of Pure Reason* to him. But because the *Critique*’s publication was delayed and did not appear until after Lambert died, this dedication sadly was

---

<sup>340</sup> “New” because its title is an implied reference to both Aristotle’s *Organon* and Francis Bacon’s *Novum Organum*.

<sup>341</sup> Kant’s letter to Lambert dated December 31, 1765 (Kant, 1997, p. 81).

not included.<sup>342</sup>

Lambert appears to have been the first philosopher to use the word ‘phenomenology’.<sup>343</sup> He considered phenomenology to be a descriptive study of the often misleading illusions [*Scheine*] that lead one to make false claims about experience. He characterized phenomenology as “the doctrine of illusion.” For Lambert, phenomenology sets for itself the goal to study the ways in which illusion can be avoided so that one can come to a recognition of truth. Beginning with its birth in Lambert’s *Neues Organon*, the term ‘phenomenology’ was associated with, in my words, a *negative, error-detecting and error-eliminating* purpose.

Kant was inspired by Lambert’s conception of phenomenology, and proposed in his letter to Lambert of September 2, 1770, what Kant called a ‘*phaenomenologia*<sup>344</sup> *generalis*’, a general phenomenology that would serve as a *defensive, preventative* study (again, my words), a study which would comprise a propaedeutic discipline to *protect* the future development of metaphysics from error-causing “contamination” from the illusions that often afflict sensory experience. In other words, this general phenomenology would determine the validity and limitations of the principles of sensibility to prevent their misapplication beyond their legitimate range, to what Kant called the ‘objects of pure reason’ studied by metaphysics. Kant called such a *phaenomenologia generalis* a ‘*negative science*’:

The most universal laws of sensibility play a deceptively large role in metaphysics, where, after all, it is merely concepts and principles of pure reason that are at issue. A quite special, though purely negative science, general phenomenology (*phaenomenologia* [sic] *generalis*), seems to me to be *pre-supposed* by metaphysics. In it the principles of sensibility, their validity and their limitations, would be determined, so that these principles could not be confusedly applied to objects of pure reason, as has heretofore almost always happened. For space and time, and the axioms for considering all things under these conditions, are, with respect to empirical knowledge and all objects of sense, very real; they are actually the conditions of all appearances and all empirical judgments. But extremely mistaken conclusions emerge if we

---

<sup>342</sup> Peters (1968, p. 453) and O’Leary (2010, p. 385).

<sup>343</sup> Lambert (1774, Vol. 1, p. 4).

<sup>344</sup> This was Kant’s misspelling in his letter. Except when quoting Kant, elsewhere in the main text I have used the corrected ‘*phaenomenologia generalis*’.



apply the basic concepts of sensibility to something that is not at all an object of sense, that is, something thought through a universal or pure concept of the understanding as a thing or substance in general, and so on. (Kant, 1997, pp. 108-109, italics added)

Kant did not, to my knowledge, use the expression ‘negative science’ again. It was for him, during the period preceding the publication of his *Critique of Pure Reason*, a proposed discipline that would precede metaphysics with the ideal aim of assuring that the illusions of sense experience do not corrupt or invalidate the results of pure reason. As he expressed this above in his letter to Lambert, such a “negative science” was to be a discipline that would determine “the principles of sensibility, their validity and their limitations ... so that these principles could not be confusedly applied to objects of pure reason.” Kant’s “negative science” was to serve as an *indispensable preliminary*, a scientifically disciplined safeguard against the subsequent introduction of error that can originate from illusory sense experience.

Negative science for Kant was—again to emphasize this—to be a “propaedeutic” discipline. The term derives from the Greek ‘*propaideuein*’, meaning “to teach *beforehand*.” This “beforehand” is important: In connection with the negative science of the metalogic of reference, this “beforehand” means that *it stands in a position of logical and transcendental priority in relation to subsequent philosophical or scientific investigations*. In this sense the metalogic of reference is conceptually more basic than is an inquiry into the transcendental foundations of knowledge. The capacity, in principle, to refer identifyingly to a set of objects must first be assured in order for it to be possible to make claims to knowledge about those objects. The focus of the metalogic of reference may, in this way, be termed ‘*maximally fundamental*’. In a theoretically fundamental sense, the task of the *Critique of Impure Reason* must precede the *Critique of Pure Reason*’s investigation of the preconditions of knowledge.

There is a significant distinction to be drawn between a *negative critique*, in Kant’s case, of the arguments of metaphysics, and his later *positive elaboration* of a theory of the principles of reason. “[M]y suggested treatment [of metaphysics] will serve a merely *negative purpose*, the avoidance of stupidity (*stultitia caruisee*), but it will prepare the way for a positive one.”<sup>345</sup> For Kant, negative science is “presupposed” by metaphysics; it must come “beforehand.”

---

<sup>345</sup> From Kant’s letter to Moses Mendelssohn, April 8, 1766 (Kant, 1997, pp. 90-91, italics added).

Kant's proposed negative science, however, is evidently not the negative science of the metalogic of reference, which is characterized by a well-defined de-projective method and has a considerably broader and more comprehensive scope of application that includes all coherent systems of reference, itself included. But in parallel with Kant's *phaenomenologia generalis*, negative science in the present study is intended to serve the purposes of a preliminary and indispensable conceptual method. Its objective is negative in both the preventative and eliminative senses: to assure that the concepts, claims, and positions formulated by philosophy, that commonplace concepts, claims, and positions which characterize everyday thought and discourse, and that the concepts, claims, and theories advanced by the positive sciences do not attempt to exceed the metalogical horizons in terms of which they can, in principle, be meaningful. But providing that assurance, a major and prevalent form of error can be averted.

### **29.2 The metalogic of reference as a theory of error analysis and correction**

Philosophical thought and the concepts, claims, and positions people embrace in everyday life have not benefitted by a rigorous and commonly accepted means to detect, avoid, and eliminate conceptual errors, and, to be sure, not of the kind identified in this study. Philosophy has largely relied upon informal logic to recognize errors made by philosophers and non-philosophers in their reasoning. Once the tools of mathematical logic had been developed, philosophers who have wished to express their arguments in the formalized languages of mathematical logic have sometimes profited by the capacity of these tools to make the structure of their reasoning more explicit, clearer, and more logically precise. As the family of logics has grown, philosophers are often able to choose a system of logic that most closely suits the tasks they wish to undertake. Formal logic often allows one to formulate proofs and to detect errors in reasoning more effectively and in a manner less vulnerable to misinterpretation and controversy. The great breadth of topics examined by philosophy has brought with it the need to "custom fit" or else to "custom design" logics in order to meet the special needs of individual topics, as we see, for example, in deontic logics, erotetic logics, relevance logics, many-valued logics, quantum logics, etc. However, in part due to the wide range of subject matters studied by philosophy, it has been challenging for philosophers to reach agreement upon a uniform set of error-detecting standards and methods.

Seldom falling within the scope of attention of philosophers, the autono-

mous specialized study of defensive and preventative means to identify and eliminate error has advanced very considerably during the past two and a half centuries. Lambert's and Kant's attempts to insulate reason from the contamination of error-causing illusions of sensation has in the positive sciences become a good deal more sophisticated and capable of disciplined application.

Not in a mere figurative analogy with the needs of philosophy, but in a directly relevant and applicable parallel, the natural sciences and mathematics have been able to incorporate in their methodologies important means of self-monitoring and avoidance of error. They have accomplished this through their adherence to strict, universally accepted standards and error-detecting methods designed to permit and encourage scientists—in keeping with the “beforehand” imperative of negative science—to determine possible sources of error *in advance* of actual experiments, observations, publication of results, the design and development of instrumentation, or the initiation of construction projects. Such sources of possible error include errors of measurement, errors of procedure, errors due to miscalculations, etc. Mathematicians expend considerable effort in proof-checking in advance of publication; computer science is evolving increasingly sophisticated automated means of error-checking computer code before it is implemented. Without such ideally rigorous methods of error-checking and error-avoidance, planes would crash more frequently, bridges would collapse, the public would be made ill by medications, mathematical advances would become questionable, etc.

To mention by way of informative illustration a specific area in which such methodical advances have been made in recent years, physics (along with many other disciplines<sup>346</sup>) has developed an entire sub-specialty devoted to error analysis and error theory.<sup>347</sup> The varieties of error that are studied include statistical or random errors, uncertainty due to the degree of precision of instruments, the domino effect of propagation of errors, and systemic errors. The latter two forms of error relate directly, as we shall see, to the concerns of the metalogic of reference as a negative science.

In physics, *systemic error*<sup>348</sup> can occur when the faulty design or construc-

---

<sup>346</sup> Specifically, two principal methods of error analysis have come to be widely employed in such areas as: safety and reliability engineering; hazardous research that includes aerospace; the nuclear power industry; the chemical, petrochemical, computer science, and pharmaceutical industries; as well as the development of safeguards to limit the potential for the unintended launching of nuclear missiles. One method is deductive in nature, *fault tree analysis*, and the other is inductive, *failure mode and effects analysis*. Both methods of error analysis are today heavily relied upon to prevent potentially disastrous failures and devastating outcomes.

<sup>347</sup> For background, see, for example, Taylor (1982) and Bevington and Robinson (2010/1969).

<sup>348</sup> By physicists frequently called ‘systematic error’.

tion of a given system or of the instrumentation that is employed to study that system leads to results that fall short of what was ideally intended. Systemic error of this kind, once identified, often makes it possible to compensate by corrections for such error and to avoid undesirable and potentially catastrophic failures and destructive consequences.

Previous chapters in this study have come to understand frames of reference as dynamic referential systems that define associated referential fields. From this standpoint, metalogical projections comprise systemic errors of a very different kind from those of concern to physicists. Unlike systemic errors in physics, the systemic errors with which we have been concerned are wholly conceptual errors, expressed by attempts to trespass beyond the metalogical horizons of systems of reference. But like systemic errors in physics, systemic errors of metalogical projection, once they have been identified, can often be corrected, either by means of de-projective modification or by elimination. From a philosophical perspective, *there is certainly no more catastrophic failure and destructive consequence for a concept, claim, or theory than to undermine its own possibility of meaning through metalogical projection.* Given the wide range of areas with which the conceptual analyses of philosophy may be concerned, there is, as in physics, a comparable need in philosophy for a uniformly agreed-upon, compelling method capable of identifying, correcting, and, if necessary, eliminating such metalogical errors of reference.

The second kind of error mentioned above, *propagation of errors*, occurs with great frequency in connection with projective concepts and in resulting claims and beliefs which incorporate those concepts, which are deceptively taken to be meaningful, and which then are often extended and increasingly proliferate in a widening diversity of domains of discourse. As we have seen amply exemplified by the varieties of metalogical projection analyzed in this book, the propagation of projective errors is both extensive and difficult to restrain.

In {12}, the negative science of the metalogic of reference was compared metaphorically with a “sieve” capable of sorting out the meaningful from the meaningless, a “filter” whose objective is to insure that only meaningful concepts, propositions, and statements can remain as the subject for subsequent analysis and potential use. The task of accomplishing this is clearly one of error analysis and correction on a purely conceptual level. To this end, the metalogic of reference comprises a theory of error analysis and correction by providing a highly general, self-validating criterion of meaning that makes it possible to identify, sometimes correct, and at other times entirely eliminate concepts, propositions, and statements that undermine themselves on the level of possible meaning. In the process, the metalogical horizons associated with

frames of reference are made evident, revealing the limits beyond which reference becomes meaningless.

### 29.3 Eliminative psychology and projective delusion

As understood in this study, the negative science of the metalogic of reference has a purely conceptual focus and conceptual objective: to study concepts, claims, and positions found in philosophy, in the conceptual vocabulary of ordinary thought, and in theories advanced by the sciences, and in that study to serve as a “sieve” or “filter” that would identify those concepts, claims, and positions that are metalogically self-undermining, correct those that can be corrected, and reject those that cannot. By realizing this central objective, the metalogic of reference helps to insure that the conceptual basis of philosophical, everyday, and scientific thought can be relied upon to be error-free of metalogical projection.

This purely conceptual focus, however, though it defines this book’s *raison d’être*, requires the extension of its scope *if* it is to include and apply effectively to the human beliefs which projective concepts, claims, and positions frequently engender. Although mention has routinely been made of the role of beliefs, it is the conceptual basis of such beliefs that has been our principal concern. In this section, I turn to consider the difficult issue posed by the human tendency to fall victim—willingly but often willfully—to projection.

Early in this book, in {2}, I engaged in a self-conscious and unashamed polemic against *baseless beliefs*—those that rest on no evidence, or that rest on no *possible* evidence, or that rest on putative evidence that would, in principle, be self-undermining. I discussed willful blindness in embracing baseless beliefs in philosophy, and, as one major contributing factor, blamed such recalcitrant blindness for the fact that philosophy has so little in the way of firmly established results to show for itself over the course of more than two thousand years.

When projective beliefs are revealed to have no possible meaning because the preconditions for their possible meaning cannot in principle be satisfied, I’ve called such beliefs ‘delusional’. Later, in {12.2}, delusional beliefs were described in greater detail, referring in part to the definition of delusion as formulated by today’s Bible of psychiatric “disorders,” the *DSM*. I then extended that definition to include the epistemological sense in which the particular variety of delusions of belief that concerns us is, first, *exempted from rational rules of admissibility* ({17.3}), and, second, tends to be accompanied by ‘*anosognosia*’ or *lack of awareness of the existence of the delusion*. In

addition, we saw that such delusions incorporate a component of *erroneous identification*, which serves as the basis for the delusional mistake that projective beliefs involve. When understood from the standpoint of reflective analysis, we saw that this component involves the deception that projective beliefs allegedly involve possible reference and hence putative meaning. In previous chapters, we have encountered this deception in a wide variety of contexts, involving a multiplicity of concepts, claims, and positions. We have come to see that the putative meaning of such concepts, claims, and positions quickly vanishes under the sharp eye of reflective metalogical scrutiny, evaporating before our eyes into meaninglessness.

Despite this reflective analysis, however strongly compelling it is for any given reader, it is unrealistic to infer that projective *beliefs* (and here I do not refer to the underlying concepts, claims, and positions which have formed our principal subject-matter) will themselves also quickly dissolve away, for such beliefs tend to be extremely recalcitrant and self-perpetuating. There is an undeniable “*compulsion to project*,” a compulsion, or drive, which centuries ago, as we saw in {14}, Kant also noticed when he asked, “[W]ho does not *feel himself compelled* [*fühlt sich nicht notgedrungen*], notwithstanding all interdictions against losing himself in transcendent ideas, to seek rest and contentment beyond all the concepts which he can vindicate by experience...?” (Kant & Beck, 1950/1783, §57, italics added). Though Kant saw this compulsion as a “transgression,” he himself repeatedly fell victim to it<sup>349</sup> (or he may even be interpreted to have intentionally embraced it<sup>350</sup>).

It is unquestionably hard to curtail the psychological drive to overstep the horizons of our frameworks of reference. Curtailing that drive, placing it in check, learning to exercise control over it, keeping it within horizon boundaries, is a psychological and not primarily a conceptual matter. In the Introduction, I mentioned that this book would contain a “mildly perceptible undercurrent of psychology.” I have deliberately given that undercurrent a very minor role, one which, in relation to the purely conceptual objectives of this study, has been entirely dispensable. In this chapter, an exception will be

---

<sup>349</sup> {6, 21.2, 22.5} and *passim*.

<sup>350</sup> In my reading, Grier (2001) offers such an interpretation of Kant: She argues that Kant developed a “doctrine of transcendental illusion” to make it possible for him to claim that, necessarily, “illusion” is involved in making the transition from sense experience to a unified understanding of reality that “goes beyond” sense experience and therefore must “hypostatize” noumena. If this is the case, then translated into the terms of the metalogic of reference, this would mean that Kant dignified what I call ‘metalogical projection’ as an essential means to acquire unified knowledge of reality. This entails that such knowledge is only possible based on concepts that are projectively meaningless. This of course stands coherent understanding of reality on its head while undermining its meaning.

made so that the psychological undercurrent is made more explicitly perceptible.

The metalogic of reference has been characterized as a “negative science”; we may also speak of a related but separable inquiry that would focus on the psychological nature of the “compulsion” or “drive” that is at the core of projective *beliefs* (but again, not at the core of projective *concepts, claims, and positions*). To complement the objectives of negative science, we may conceive of a set of objectives which set as their task the study and restraint of that psychological compulsion to transgress metalogical horizons. Let us call this the ‘*eliminative psychology* of the metalogic of reference’.

It makes sense to think of attempts to restrain and correct the psychological drive to trespass beyond the boundaries of metalogical horizons as a form of *eliminative psychology*. Its aim is explicitly purgative: to *recognize* and when possible to *eliminate* the psychological propensity to invest belief in that which cannot, in principle, be meaningful. Alternatively conceived, its aim is to *remove the psychological obstacles to de-projection*.

If would be a great mistake to suggest that such a task, given the present human psychological constitution, has much hope of widespread success. The human horizon-transcending compulsion is psychologically deeply rooted; it is, from all the anthropological evidence, a primitive response. When physical conditions have been harsh, dangerous, and life-threatening, the psychological drive to wish that things were otherwise has, throughout human history, taken many forms as we see in the endlessly spun tangle of myths, mythologies, and religions. Modern human beings are in this respect very likely not much different from their distant ancestors. Under duress and high anxiety, the compulsion to reach for reassurance, succor, comfort, and security quickly becomes a dominant force in many people. When drowning, the struggling man will often reach for salvation that lies beyond his grasp. We see the same reaching for what lies beyond one’s grasp in hope for what is not, in faith in what is not, in love for what is not. This phenomenon, in all its sizes and shapes, occurs everyday when people encounter hardships and challenges they feel they cannot cope with on their own.

It is also a response that affects philosophers when they encounter many conceptual problems whose solutions seem to “demand” that philosophers “go beyond” available experience, and not only this, but whose solutions appear to require that philosophers “go beyond” all that to which the frameworks of reference of those problems make it *possible* to refer. The “duress and anxiety” provoked by imminent physical danger have their intellectual counterparts when they take the form of philosophical problems that present conceptually as well as emotionally demanding challenges. When viewed

from the standpoint of the metalogic of reference, a common philosophical response to such problems has been to engage in conceptual *over-reaching* that derives from the epistemologically naive human belief that “reality must be more than we know” and the class of beliefs related to this. We have seen that such horizon-transgressing beliefs are typically asserted categorically, believed with determination, often unhesitatingly and cavalierly dismissing contrary evidence, as if the force with which they are adamantly asserted has, in itself, persuasive power—which, for many believers, it unfortunately does. In the case of metalogical projections, one is reminded of the dictum: “*credo, quia absurdum*” (Freud 1930/1952, p. 786)—I believe *because* it is absurd — a maxim that Freud of course rejected. When a conceptual projection and belief in that projection meet, as noted in {20.10}, the projective conceptual drive and the psychological compulsion then intertwine and mutually reinforce one another.

Is there a solution to what Kant called this human ‘*natural tendency to transgress*’ (Kant, 1929, A642, B670, p. 532)? He did not reject that tendency, nor did he commit himself in his philosophical system to avoid it, but still he did raise the question. I think there is a solution, but it is, as one would expect, predominantly psychological in nature (and perhaps even to a degree biological). A discussion of that solution will not be included here, but will be found in Appendix II.

### 29.4 Obstacles to attempts to de-project belief

Beware how you kill a thought that is new to you.

– George Bernard Shaw (1934, p. 183)

[T]wo nations professing incompatible philosophies put them to the test of force. Philosophically, it was idiotic, for while opinions were arguable, *convictions needed shooting to be cured*; and the struggle could end only when the supporters of the one immaterial principle had no more means of resistance against the supporters of the other.

– T. E. Lawrence, AKA “Lawrence of Arabia”  
(1935, p. 190, italics added)

Revisionary attempts to change beliefs can, at the very least, be a struggle, and they can, in the extreme, prove in practice to be impossible. There are, to



be sure, many kinds of beliefs, held with different degrees of commitment, and expressing a wide range of motivation. To offer a detailed psychology of belief would require an entire book and is therefore not the objective here. However, much of this study has been occupied with the revisionary work of identifying and correcting concepts that undermine themselves on a metalogical level, concepts that frequently are at the core of many beliefs. The negative science that wishes to accomplish this revisionary work would be rendered ineffectual on an applied, practical, and not purely conceptual level if it did not extend its revisionary efforts to include projective beliefs. There are specific obstacles to such an extension, obstacles for which human psychology is unquestionably responsible. In this and the next section, I wish to give briefly and only in outline a description of some of the major obstacles that stand in the way of our breaking free from the by-now-familiar compulsion to engage in metalogical projection. Identifying these obstacles will enable us to distinguish several basic principles that describe how human psychology routinely serves to constrain the minds of believers so they will reject attempts—frequently in a very forceful way—to revise their projective beliefs. These fundamental psychological principles of projective belief show how such beliefs maintain and preserve their often tenacious and unyielding hold on the minds of their believers, and succeed in propagating themselves. Our specific interest, then, will be the psychological principles that underlie resistance to the elimination of metalogically projective beliefs.

Previous chapters have described how frameworks of reference function as dynamic systems of reference which satisfy the preconditions of reference required for concepts, claims, or positions, in principle, to possess possible meaning. Holders of beliefs that employ such concepts or assert such claims or positions must presuppose the reference frames necessary for those concepts, claims, or positions to have possible meaning; and, to be sure, the holders of beliefs believe their beliefs are meaningful. In other words, holding putatively meaningful beliefs necessarily relies upon the referential framework(s) in terms of which they are maintained.

Years ago, philosopher William Todd (in a book unrelated to our subject here) proposed that to hold a belief is “to feel the emotion of surprise upon discovering that [the belief] is not the case” (Todd, 1968, p. 310). I do not think this suggestion is generally true—it is certainly not true, as we shall see, of strongly held beliefs; nevertheless, the very reaction of surprise on the part of believers deserves some attention: Committed believers in an ideology, for example, are never subject to such “surprise” for the simple reason that they are, by virtue of the articles of their rigidly held dogma, “immune” to the discovery that their beliefs are not true. But surprise can certainly be aroused

when an ideologue is confronted by someone who does not share in his or her ideology and questions or opposes it, and then surprise may quickly give way to the believer's cognitive and emotional "dissonance," experienced in the form of shock, frustration, then perhaps anger, and perhaps verbal or physical aggression.

This latter kind of surprise is instructive: It is an expression of the psychological principle that functions to "immunize" believers from incompatible beliefs, preventing believers from breaking free from the frameworks of reference which their beliefs presuppose. This phenomenon is very familiar to most of us: it is the *immunization of belief* to contrary evidence. It is a phenomenon previously considered in more detail in {1.2}.

There are four varieties of belief and of holders of belief that we may distinguish, two of which are relevant to the phenomenon of immunization of belief:

	<b>Beliefs are based on evidence</b>	<b>Believer is open to revisions of his/her beliefs in the light of new evidence</b>
1	yes	yes
2	yes	no
3	no	yes
4	no	no

**Table 29.1. Openness of beliefs to revision in light of evidence**

Cases 1 and 3 identify believers who are sufficiently open-minded and flexible in their adherence to their present beliefs to revise them when presented with new evidence. Cases 2 and 4 relate to believers who are closed to new evidence: Both of these groups of believers will hold fast to their beliefs, whatever may be the new evidence with which they are presented. They make up the class of believers whom empirical science is unable to reach and convince.

Let us now consider four analogous yet fundamentally different varieties of belief and of holders of belief:

	<b>Beliefs are free of metalogical projection</b>	<b>Believer would be open to revisions of his/her beliefs were any of these to be projective, and were shown to the believer in a strongly compelling manner that they are metalogically self-undermining</b>
1'	yes	yes
2'	yes	no
3'	no	yes
4'	no	no

**Table 29.2. Openness of projective beliefs to revision in response to strongly compelling rational demonstration**

Again cases 1' and 3' identify open, flexible thinkers, whereas 2' and 4' relate to individuals who will persist in their beliefs whether or not those beliefs are projective even when these beliefs are shown to them in a strongly compelling manner to be projective. Cases 2' and 4', which concern us, are characterized by a particular form of immunization of belief to metalogical argument. What is the special nature of the psychology of believers in these two cases?

At this point, I remind the reader of a group of observations concerning the psychology of belief described in {1}. There, in connection with the psychology of philosophers, I remarked on the need among philosophers to possess the ability to be receptive to and to comply with the requirements of reason. This is an ability defined by a receptivity sufficiently accommodating and resilient that an individual is open to changing his or her convictions in accordance with the mandates of reason, combined with a determination and capacity to change those convictions when required by reason. In this connection, I introduced the "rational bridge problem" to bring attention to bear on the frequent lack of genuinely felt convictions as well as corresponding conduct which ideally are the goal of conclusions reached by rational means. When a person lacks an effective bridge between rationality and conviction, I referred to this condition as a "bifurcation" of the individual's mind. When a believer possesses no "rational bridge" even when rationally persuaded, he/she will not develop the conviction necessary to bring about a fundamental change in his/her beliefs.

Implicit in the observations described in {1} are the following elementary psychological principles that not only frequently govern and limit the effec-

tiveness of communications among philosophers, but apply generally to believers who hold strongly and uncompromisingly to any set of beliefs. These principles serve basic, sometimes defensive and at other times offensive, functions that immunize a system of beliefs from change.

- ◆ *ideological self-encapsulation*: monadically walling off one's preferred system of beliefs from challenge
- ◆ *recalcitrance* (or imperviousness), which accompanies ideological self-encapsulation: an unwillingness or inability to make real in one's own mind a frame of reference perceived to be fundamentally different and alien; this is an inability to comprehend an alternative and fundamentally dissimilar conceptual system, so to speak, from within
- ◆ *distortion*: when one's own preferred system of beliefs is subjected to criticism, this is a disposition to react by misconstruing, misstating, or misapplying the critic's position, by denial and rejection, and by shifting the grounds of debate

These principles apply to the two cases 2' and 4', i.e., to individuals who will persist in their projective beliefs whether or not those beliefs are demonstrated to them to be metalogically self-undermining.

However, unlike familiar instances in which a set of beliefs is challenged by an opponent's differing set of beliefs, here the challenge to a set of projective beliefs is of a different kind; it is in an important sense *intrinsic*: Here, a strongly compelling demonstration which shows believers in a set of projective beliefs that those beliefs are metalogically self-undermining does not make recourse to a separate opposing system of beliefs. Rather, such a demonstration shows that the very possibility that the beliefs in question should have meaning is ruled out by referential preconditions undermined by those beliefs themselves. Suppose, as an example, a believer believes in the putatively meaningful notion of "an autonomous reality, existing independently of all reference frames." A reflective analysis will show that the very referential preconditions that would need to be met in order for that notion to have possible meaning are conditions the satisfaction of which the notion itself precludes.

To re-state the question with which we're concerned: What characterizes the psychology of projective believers who fail to be convinced by such a demonstration, a demonstration that they must revise their beliefs on pain of undermining the very possibility that their beliefs should possess meaning?

We can give an obvious, correct, but overly simplified answer, that such believers are not rational beings and/or they are committed to ideological self-encapsulation, recalcitrance, and distortion. This answer is likely to be correct but it does not go far enough.

It does not go far enough when we find that one belief after another in a given system of belief is projective, and also find the same pattern of persistent belief in projections in one system of belief after another. When we find that essentially the same group of projective notions are believed to be meaningful by believers who hold often divergent systems of belief, we should be moved to ask for a more comprehensive answer. We begin to suspect that more is involved in addition to potential irrationality and the role of psychological principles that bring about the immunization of beliefs to revision.

In previous chapters, we have analyzed a wide array of projective notions. Our analyses have intentionally remained on a purely conceptual level—Independently of and unconcerned with the pragmatical application of the results we have reached, without concern for their application to the individual thought processes of people whose beliefs are based on those projective notions. If we extend our scope of interest in this pragmatical direction, we run not only into the defensive and offensive walls erected by the principles that immunize systems of belief from challenge and change, but we must confront the persistent “natural human tendency” to wish to transgress beyond horizons of possibility and meaning. To accept the simple answer above leaves us without a psychological understanding of this natural human tendency, and therefore without a complete understanding of the persistence of believers to hold fast to beliefs that are metalogically self-undermining. It fails to inform us why anyone would continue to hold such beliefs who has been shown that his or her beliefs are projective by means of a strongly compelling demonstration—one that cannot *not* be accepted without undermining its possible meaning. Let us look at the question from a different and potentially more promising perspective.

### **29.5 From conceptual therapy to disorders of thought: The human will to reach beyond its grasp**

In the Introduction and in {13.1}, I described an approach that I’ve called ‘conceptual therapy’ to refer to a form of conceptual analysis that has a “therapeutic” goal: to identify, correct, replace, and avoid metalogically self-undermining concepts. This endeavor, as has been emphasized, is exclusively conceptually centered and conceptually applied. It is a “therapy” for malfunctioning concepts as distinguished from a therapy intended to assist people in

eliminating and avoiding their own acceptance and use of such concepts. If we decide to attempt in this way to extend such a conceptual therapy to persons, that effort can easily run aground on the shoals of deficient reasoning and individual belief system resistance, defensiveness, and opposition. The practical implementation of conceptual therapy is a task of education or skill-based training. Unlike the limitation of the metalogic of reference to a therapy for concepts, the educational objective is not a science, but rather involves a discipline that can potentially be developed, cultivated, and sharpened, but one that is nonetheless a discipline subject to the vagaries and unreliability of all efforts to educate, train, and improve any group of skills.

If we shift our discussion of a therapy intended for concepts to a corresponding therapy intended to remedy how well people think, it may come as an affront to some readers that there may be such things as “*disorders of thought*” that can stand very effectively in the way of a successful transition from conceptual therapy to a therapy applied to the competence and manner in which people think. It may therefore be appropriate to describe briefly the basis for the claim that disorders of thought are real, and to refer to previous research and publications in which this claim is supported. It would, however, take us too far afield to include a detailed discussion of past research in psychology and psychiatry having to do with the general category of thought disorders.<sup>351</sup> To keep this discussion brief, I will mention only work that relates to the attempt here to understand the natural human compulsion to engage in projective thinking.

One of the earliest publications to direct attention to what today have come to be known as “disorders of thought” was Kant’s small book, *Anthropology from a Pragmatic Point of View* (Kant, 1974/1797). It is considered to be Kant’s work most relevant to today’s now-independent discipline of psychology. His book provides a short catalog of a number of mental dysfunctions or illnesses. Book I of the *Anthropology* is titled “Deficiencies and Diseases of the Soul with Respect to Its Cognitive Power,” a clear indication that Kant intended to consider *cognitive* disorders or malfunctions, and less what today are commonly judged to comprise “mental disorders” which generally involve dominant emotional components. Among the disorders that he lists are two which we would likely consider to be more intellectual and cognitive in nature; they relate to how the mind makes use of concepts. One of these he called ‘*insania*’, in which the mind mistakes analogies “for concepts of things similar to each other,” and the other he called ‘*vesania*’, in which

---

<sup>351</sup> One of the best concise retrospective summaries of this research may be found in Andreasen (1979). A more recent, detailed account and discussion of supporting evidence will be found in Bartlett (2005, 2011).

there is a breach with reason; an individual afflicted by *vesania* “fancies that he conceives the inconceivable” (Kant, 1974/1797, pp. 84-85). The latter involves what today we would call ‘delusional thinking’. It represents one of the first times a flawed pattern of thinking was identified as a disorder of thought, or what Kant called a cognitive ‘deficiency’ or ‘disease’.

We should immediately recognize the relevance of such a disorder of thought to the earlier discussions in this study of Kant’s notion of “compulsion”<sup>352</sup> expressed in the human propensity to overstep boundaries—in Kant, these are the boundaries of transcendental inquiry. He claimed that it is a compulsion that leads to “delusions” and “deceptions” (*Wahn* and *Blendwerke*).<sup>353</sup> In his *Anthropology*, Kant came close connecting the “*natural tendency to transgress*” with a psychologically based propensity, one that can justifiably be judged to be a cognitive disorder.<sup>354</sup> If we can shed light on the nature of this form of “delusional thinking,” we will come closer to a psychological understanding of the pervasive propensity of people to invest extremely resistant belief in metalogically projective notions.

As far as I have been able to determine, after Kant’s *Anthropology*, the next step in an attempt to throw informative light on our species’ “boundary-overstepping propensity” was taken beginning in the early 1960s when the present author began to develop “conceptual therapy.” By the mid-1960s, I had provided an account according to which there exist real, non-metaphorical pathologies of a conceptual kind—“conceptual pathologies”—which significantly distort and undermine human thought. This work led to my doctoral dissertation (Bartlett, 1970), followed by a series of related publications.<sup>355</sup>

Related to this research, in my university teaching I began in 1971 to develop a skill-based approach to train students in the skills of logically based epistemological analysis, and specifically the skills of de-projective analysis.<sup>356</sup> I was drawn by the possibility that a transition could be made from conceptual therapy to the development of related cognitive skills. In this effort, I

---

<sup>352</sup> See {14} and {20.11}.

<sup>353</sup> See {14, note 209}.

<sup>354</sup> Kant evidently believed that the treatment of the cognitive disorders he catalogued should fall within the province of philosophy. However, the disciplines of psychology and psychiatry, which did not exist in Kant’s time, today judge such disorders to fall within their disciplinary territories. The author, nonetheless, sides with Kant on this issue, since the teaching of skills in thinking according to the dictates of reason continues to be a task delegated to teachers of philosophy and logic, not to psychologists and psychiatrists. One cannot realistically imagine a class in either of the two latter disciplines designed to train students in the skills of de-projective analysis.

<sup>355</sup> See References.

<sup>356</sup> Cf. Bartlett (1973, 1976-77, 1978-79, 1983a, 1983b).

learned that the “natural tendency to transgress” found in my students was enormously resistant to change. This, in turn, led me to undertake research from the perspective of clinical psychology relating to this prevalent human disposition; I continued this study over a period of decades.<sup>357</sup>

In 1979, American psychiatrist and neuroscientist Nancy Andreasen offered a detailed classification of 18 varieties of thought disorders that are frequently encountered in psychiatric patients, some of which are also found in the psychologically normal population. She observed that “thought disorder is a continuous rather than discrete phenomenon and ... it shades gradually into normality” (Andreasen, 1979, p. 1325). None of the thought disorders she identified, however, relate directly to metalogically projective thinking. Two that might be said to be indirectly associated are “derailment” (“Things may be said [by a patient having this disorder] in juxtaposition that lack a meaningful relationship, or the patient may shift idiosyncratically from one frame of reference to another.”) and “illogicality” (“[Q]uite common [also] among nonpatients.... Illogicality may either lead to or result from delusional beliefs.”) (p. 1319).

In {12.2}, I described more recent research that has come after Andreasen’s classification; this work concerns a class of disorders of thought that have come to be called ‘delusions of misidentification’. Also in {12.2}, I proposed a closely related variety, delusions of meaningfulness. That variety now becomes relevant to the question at hand, What is the special nature of the psychology of believers in the two cases, 2’ and 4’ in Table 29.2?

We recall that such delusions of meaningfulness involve two components: a delusional belief associated with an erroneous identification. The belief is erroneous, is rigidly adhered to, is exempted from criteria of rationality, and possesses an immunity to revision even in the face of incontrovertible evidence to the contrary—in the case of projective beliefs, that evidence is provided through self-validating rational demonstration. When it has been shown, in a manner that is strongly compelling, that a projective belief undermines its own possibility of meaning, the notion that the belief has meaning must, indisputably, be in error. The delusion of the putative meaningfulness of the projective belief should then—one would expect—be dispelled. But in the two cases that concern us, 2’ and 4’, victims of projective delusions persist in their projective beliefs. That they do this can be explained in individual instances, as we have said, by appealing to psychological explanations of the immunity to revision of many strongly held beliefs. But since projective beliefs take a multitude of forms, and occur with such frequency, we wish for a

---

<sup>357</sup> Certain of the results of this research are described in Appendix II, “Epistemological Intelligence.”



more comprehensive psychological account of the delusion of meaningfulness that is involved.

### 29.6 The psychology of projective belief

“Faith” means the will to avoid knowing what is true.

– Friedrich Nietzsche (1918/1895, p. 71)

The delusion of meaningfulness expressed in projective beliefs is characterized by the previously described properties of rigidly maintained beliefs, but, in addition, this special variety of delusion has several properties which, together, are especially responsible for its prevalence among believers, its persistence in the face of rational proof of its self-undermining nature, its ability to propagate, and its near-ineradicability.

First, given its great pervasiveness among people, the delusion of meaningfulness that is associated with projective beliefs is clearly neither unusual nor uncommon. Many readers may be unfamiliar with the extension of such terms as ‘disorder’ and ‘pathology’ in a manner that can at times include psychologically normal individuals. However, there is a considerable body of evidence which shows that, when conditions are right, many psychologically normal people exhibit dispositions to behave and to think in destructive ways which meet the conditions of real, non-metaphorical pathologies. This is a substantial and complex topic of itself, and cannot be treated here. Readers wishing for the theoretical basis and empirical evidence for this claim are referred elsewhere.<sup>358</sup> For our purposes here it is sufficient merely to recognize that the delusion of the meaningfulness of projective beliefs is quite common, and, to the extent that it comprises a metalogically self-undermining, short-circuiting of possible meaning, this form of delusion may appropriately be considered, when assessed from the standpoint of coherent rational functioning, to comprise a form of pathology, indeed of “conceptual pathology.” This is the first defining property of such delusions of meaningfulness: their comparative *normality* among people.

One of the principal difficulties which individuals experience if they are pressed to critique their own conceptual system—to put their own system of concepts literally in question—is the experience they have of *the absence of any other way of thinking*. When confronted by a differing conceptual system that is unfamiliar, a system of concepts perhaps not only alien but challenging

---

<sup>358</sup> See Bartlett (2005, 2011, 2013), among other related publications by the author.

to their own, many people experience an unsettling inability to comprehend. In an anthropological study of the Azande people of Africa, E. E. Evans-Pritchard discovered precisely this. “They reason excellently in the idiom of their beliefs, but they cannot reason outside, or against, their beliefs because they have no other idiom in which to express their thoughts” (quoted in Polányi, 1952, p. 221).

There can be no doubt that, from the viewpoint of people who accept and endorse projective beliefs, and for whom such beliefs may form much of the core of their conceptual system, any attempt to place their system of concepts and beliefs radically in question will not only be experienced as challenging, but as placing an excessive intellectual demand upon them exactly because they lack any alternative set of concepts from the standpoint of which they may gain sufficient distance to appraise their own. This is the second property of delusions of meaningfulness: the absence of any other way of thinking, *the lack of any conceptual alternative*.

As we have seen in some detail in this study, the analysis of metalogical projections demands reflective ability and some degree of proficiency in logical thinking. We’ve also seen that a major contributor to the immunity of beliefs to revision is the absence of a “rational bridge” that would join rationality with conviction. Certainly those who are unable to be persuaded to change their beliefs as a result of a strongly compelling rational demonstration appear to lack such a “bridge,” but there can be a more obvious reason: They may, in the words of philosopher Max O. Hocutt, be “logically obtuse”—that is, those “who do not always deduce (come to know) what obviously follows from what they do know” (Hocutt, 1972, p. 435). In more general terms, believers who persist in projective beliefs, even after those beliefs have logically been shown to be without possible meaning, may very well lack competence in logic. *Lack of both reflective ability and competence in logic* is a third defining property of delusions of meaningfulness.<sup>359</sup>

A fourth defining property of such delusions is the *emotional gratification* they frequently provide. Hope and faith, for example, often involve projective thinking that seeks to transcend the confines of immediate experience, yearning for what is believed to lie beyond whatever physical or mental conditions with which a person may be struggling, for that which is believed to possess an autonomous reality, lying beyond reach. When they posit a reassuring reality that lies beyond one’s grasp, hope and faith offer emotional support and security. As noted in {29.3}, under conditions of duress and high anxiety, this

---

<sup>359</sup> As described in Appendix II, more than these competencies is required for effective de-projective analysis. For simplicity, only reflective ability and the ability to think logically are mentioned here.

is especially true. The emotional gratification received from delusions of meaningfulness is a property that will strongly obstruct efforts to question and eliminate them.

There is a fifth and final characteristic of projective delusions of meaningfulness which we need to identify. Such projective delusions are highly responsive to the human need to conform with the beliefs of others. When many people in a group invest their credence in metalogically projective beliefs, individuals experience pressure to conform to the mentality of the herd. When many of the concepts which form a group's vocabulary of thought are self-undermining on the level of possible meaning, there is undeniable and difficult-to-resist coercive force to comply in conformity with the group's manner of thinking. We therefore must recognize the strong role of *conceptual conformity* in perpetuating delusions of meaningfulness.

To summarize:

We have described a number of psychologically based principles that help us to understand the resistance of believers to revise their beliefs. In {29.4} our interest was in beliefs considered in general; we noted that strongly held beliefs can be immune to revision, and came to see that a rational bridge cannot simply be assumed to exist in all people. The role of ideological self-encapsulation, recalcitrance, and distortion was described.

In {29.5} we made a transition from the dominant conceptually centered concerns of this study to their possible practical application to people; in making this shift, we noted Kant's early attempt to recognize and classify disorders of thought, paying special attention to his recognition of the role of delusion in overstepping conceptual boundaries.

In the present section we have considered the special variety of delusion of meaningfulness that is found in projective belief, and identified the following five psychologically based properties that contribute strongly to the unwillingness and/or inability of people to dispel their projective delusions of meaningfulness:

- ◆ the comparative normality among people of projective beliefs
- ◆ the absence in many people of any other way of thinking, that is, their lack of any conceptual alternative
- ◆ lack of both reflective ability and competence in logic
- ◆ emotional gratification received from projective beliefs
- ◆ the need for conceptual conformity

These five properties, taken together, help us to understand more clearly the prevalence of projective belief, its persistence despite rational proof of its self-undermining nature, its ease of propagation, and its near-intractability.

In this light, we may identify the *cognitive disorder of projective delusion*. It is a disorder of thought, a real, non-metaphorical cognitive dysfunction; it is a cognitive disorder which Kant came close to anticipating in the form of *vesania*, which, as we've noted, he recognized to comprise a cognitive "deficiency and disease." Disorders of thought are now routinely acknowledged in psychiatry and clinical psychology. They are a now-familiar phenomenon which, as Andreasen noted, "shades gradually into normality."

There is a clear constructive purpose served by classifying a dysfunctional pattern of thinking as a "disorder" or "pathology." Doing this is not polemical. When we identify a phenomenon on a biological level to constitute a pathology, it is in order to recognize that it is harmful and to direct attention to its potential treatment. When a pattern of thought is similarly classified, it is also to direct attention to its harmful character and to the possibility that it may be remedied.

### 29.7 "Beyond belief"

[Y]ou have no right to believe what you cannot prove....

– Somerset Maugham (1950, p. 656)

The title of this chapter, the title also of this section, has two main intended meanings: First is the exclamation we often express when a situation or behavior becomes so utterly exasperating, unconscionable, or astounding that it is "beyond belief!" I'll call this the '*exasperation reaction*'.

The second meaning has two combined components: one part descriptive and the other part imperative. In relation to the first part, the phrase 'beyond belief' may be thought to refer to those "epistemic states" which lie beyond the uncertainties, unreliabilities, or deceptiveness of belief, that is, cognitive states which, in contrast, are characterized by properties which place them higher on the scale of certainty, reliability, and truth—cognitive states that go "beyond belief." This is the descriptive component. Then there is the command implicit in Maugham's overstated attempt to give unfounded belief a bad name: the imperative that one must go "beyond belief." I'll call the second meaning of 'beyond belief', combining the descriptive and the imperative components, the '*directive prescription*'. It is the prescription that one must turn one's back on unfounded beliefs and go beyond them in order

to reach a more certain, reliable, and true understanding.

In this book, a good deal of energy, time, and space has been devoted to a purely theoretical analysis that has provided a partial inventory of purportedly meaningful but metalogically self-undermining concepts and claims which are widely accepted and endorsed. We have seen how these putatively meaningful but projective concepts and claims can and are employed by beliefs in which those projective concepts and claims play a central role.

From the vantage point that one attains in complying with the rules of rational admissibility described in {17}, by now—so it is hoped—readers will be in a position to express the “exasperation reaction” when they encounter fundamentally projective concepts, claims, positions, and theories. It is hoped that such readers will be in a position both to exclaim that the intuitive, habitual, usually unquestioning acceptance and adherence to such metalogically self-undermining concepts merit intellectual impatience and dismissal, and to respond with justified impatience that projective concepts are simply so absurd as to be “beyond belief!”

More than this, it is hoped that readers will, by this point, be in a position intellectually to recognize the full force of the imperative to go “beyond belief” in all fields of study that depend upon and respect conceptual integrity and coherence, and in doing this, to endorse and observe the “directive prescription” to relinquish and then go beyond projective beliefs.

This book was written with the problems, interests, and conceptual biases of philosophers primarily in mind. But, as has been pointed out on numerous occasions, the central role played by projective concepts and beliefs is common and proliferates wholly unchecked; projective concepts and beliefs appear and re-appear in discipline after discipline—projective concepts and projective thinking are by no means the monopoly of philosophy. And so it is to be hoped that after these many pages, assiduous, reflective, and rationally well-endowed readers will have been able to “exorcize” from their conceptual vocabularies, as well as from the beliefs in which they invest conviction, those concepts, claims, positions, and theories that are metalogically self-undermining, by correcting those that are capable of coherent rational correction, and by willingly relinquishing those for which this cannot be accomplished.

Bertrand Russell once commented in a televised interview:

“When you are studying any matter or considering any philosophy, ask yourself *only* ‘what are the facts, and what is the truth that the facts bear out.’ Never let yourself be diverted ...

by what you wish to believe.... [L]ook *only* and *surely* at *what are the facts.*” (Russell, 1959)

From the standpoint of the present study, it should be clear that “the facts” are not only those that are empirical, but facts of many kinds are numbered among them. Of particular significance in this book, these are facts about reference, about possibility, about meaning, and about the horizons of these. If read with all such facts in view, Russell’s admonition can be taken to heart in our efforts to go “beyond belief,” leaving the deception and delusions of projective concepts and beliefs behind.

## *Critique of Impure Reason:* **Its Results in Retrospect**

[I]f we are bold we shall gain one of two advantages; either we shall find what we seek, or we shall be less likely to think that we know what we do not know—in either case we shall be richly rewarded.

– Plato (1952, p. 536)

[T]here is certainly no more catastrophic failure and destructive consequence for a concept, claim, or theory than to undermine its own possibility of meaning through metalogical projection. {29.2}

### **30.1 The negative science of the *Critique of Impure Reason***

The chief interest in this study has been to develop a self-validating method by means of which we are able to identify, correct, and, when correction is not possible, to reject a large and influential group of concepts and claims that conflict with their own preconditions of possible reference and meaning. I've called this by-now-familiar method 'de-projection'. The method of de-projection was developed from the standpoint of a self-referential metatheory formulated on the level of maximum theoretical generality; I have called that metatheory 'the metalogic of reference'. Its central purpose has been to enable us to recognize the boundaries of what is referentially forbidden, i.e., the limits beyond which reference necessarily becomes meaningless. I have called these boundaries 'metalogical horizons'. When these horizons are transgressed, we fall victims—very often unwittingly and habitually—to a particular variety of widespread conceptual error that is inherently deceptive. I have called this form of error 'metalogical projection', or simply 'projection'. Projections lead to delusions of meaning that lie at the very heart of many major traditional philosophical problems and questions. By making such horizons explicit, the domain of possible sense is determined.

I have characterized this undertaking as a “negative science” (12, 13.5, 15.2, 28, 29.1}). It meets the requirements of a science by comprising a disciplined, methodologically strict approach that is defined by a clearly formulated methodology and a set of rigorous standards of demonstration which, employed together, can serve as a shared unitary, reflective, analytical framework of evaluation and judgment (29.1}). Such a negative science comprises a rigorous approach to the error detection of projections, their analysis, and their correction (29.2}). Its proper domain of application is conceptual, not performative or pragmatical. Nonetheless, its purely conceptual objective is capable of being broadened and complemented by means of an applied, pragmatical, eliminative psychology whose aim is to recognize, and, when possible, to progress beyond the human psychological propensity to believe in that which cannot, in principle, be meaningful (29.3}). This pragmatical extension falls outside of the purpose of this work; some brief observations and suggestions relating to such an extension were included in the previous chapter (29.3–29.6}).

### 30.2 The positive value of negative results

We can regard a science of the mere examination of pure reason, of its sources and limits, as the propaedeutic to the system of pure reason. As such, it should be called a critique, not a doctrine, of pure reason. Its utility, in speculation, ought properly to be only *negative*, not to extend, but only to clarify our reason, and keep it free from errors which is already a very great gain.... Its purpose is not to extend knowledge, but only to correct it.... Such a critique is therefore a preparation, so far as may be possible, for an organon....

[I]f what is designed be a critique to guard against errors of judgment (*lapsus iudicii*) in the employment of the few pure concepts of understanding that we possess, the task, *merely negative as its advantages must then be*, is one to which philosophy is called upon to devote all its resources of acuteness and penetration.

– Immanuel Kant (1965/1929, A11-A12, B25-25, pp. 58-59, and A135, B134, p. 179, italics added)



The negative results we have reached in previous chapters concern a very wide range and variety of concepts, claims, positions, and theories that have been shown to be projective. What is the positive value served by such a demonstration? Certainly the elimination of conceptual error, especially a major form of error that is both widespread and generally unrecognized, constitutes a positive service to coherent thought. It will be useful briefly to take stock of the many projective concepts and claims which in this book we have identified, and have either corrected or replaced with de-projected revisions, or else have rejected entirely:

1. projections of discovery {19.1–19.2}
2. projections of invention {19.1–19.2}
3. projections of the finitude of knowledge {20.2}
4. projections of the incompleteness of knowledge {20.3}
5. projections of the unlimitedness of our ignorance {20.4}
6. projections of thinking beyond the limits of thought {20.5}
7. projections of expressing the inexpressible {20.6}
8. projections of unknown truths {20.7–20.8}
9. projections of unanswerable questions {20.9}
10. projections of the external world {21.1}
11. projections of things-in-themselves {21.2}.
12. projections of other minds {21.3}
13. projections of other minds as things-in-themselves {21.3}
14. projections of general continuity {21.4}
15. projections of realism {21.5}
16. projections of idealism {21.5}
17. projections of the past {22.2.1}
18. projections of time-flow {22.2.2}
19. projections of the future {22.2.3}
20. projections of absolute time {22.2.4}
21. projections of temporal constitutive subjective activity {22.2.5}
22. projections of absolute space {22.4.1}
23. projections of spatial constitutive subjective activity {22.4.2}
24. projections of temporal or spatial continuity {22.5}

25. projections of space-time {22.6}
26. projections of counterfactual causation {23.1.6}
27. projections of genetic causation (alternatively, projections of causal agency) {23.1.7}
28. projections of framework-transcending causality {23.1.8}
29. projections of hidden determinants {23.2}
30. projections of reflection {24.1–24.4}
31. projections that thinking entails a thinker {24.1–24.4}
32. projections of the self as an existing entity {24.5}
33. projections of the self as center of experience {24.6}
34. projections of the self as bearer or owner of its states {24.7}
35. projections of mental faculties {24.8}
36. projections of agency {24.9}
37. projections of spectator consciousness {24.10}
38. projections of consciousness as a container {24.11}

This is by no means an exhaustive list of the seemingly inexhaustible proliferation of ways in which concepts, claims, positions, and theories may undermine themselves on a metalogical level. As we have seen in all of the above instances, the human projective urge recognizes no bounds in formulating and embracing horizon-transgressing notions, while, in the process, that urge frequently invests willful and recalcitrant belief in the delusions of meaning that result.

In connection with many of the above-listed projections, de-projective corrections or replacements have been described which salvage from each putatively meaningful concept, claim, position, or theory what can be saved, and dispensing with those that cannot. The result of this process can be characterized in the following semi-metaphorical way:

### **30.3 Delineating reality in silhouette**

A silhouette is commonly thought of as a representation of a physical body, very often a person's face, that is crafted by cutting a dark material in a shape that preserves a likeness of the original. A silhouette circumscribes an outline of the original object by cutting away that which does not belong to it. In this sense, the process of cutting out a silhouette may be thought of as a negative process, eliminating what is unnecessary and therefore dispensable. A well-

made silhouette serves as a true likeness of the original, whose identifying features have been captured by the silhouette artist.

Let us leave the silhouette metaphor behind, and speak instead not of “representing” reality, but of understanding it. We recall that to *be* is to be an object of reference relative to an appropriate framework—that for something to *be* is a function of the coordinative relation of object of reference in relation to an appropriate reference frame in terms of which it can be identified, whether vaguely or precisely, or according to rule ( $\{18.3\}$ ). “Reality”—as understood in these terms from the standpoint of the metalogic of reference as the totality of possible objects of reference in the ontologically open sense described in  $\{18\}$ —is functionally defined in framework-relative terms. To ask whether or to claim that reality has—or has not—an existence autonomous of the systems of reference a function of which reference to reality is possible, is to ask or claim that which can have no possible meaning.

When many of the concepts in terms of which we understand reality are projectively deceptive—fostering claims, positions, and the formulation of theories that incorporate those projections—it is possible to come to a coherent, intelligent, and meaningful understanding of reality *negatively* through a disciplined process that eliminates projective error. As we come to recognize, correct, and sometimes entirely reject one projection after another, the understanding we come to *of that which remains* is purged of the human propensity to transgress horizons of reference. When we then become aware of and comply with the inevitable limitations that come with the adoption of any rationally coherent standpoint, much that was projectively believed to be possible and to possess purported meaning must be relinquished. It is this conceptual process of letting go of projectively delusional concepts and claims, and of correcting or rejecting the positions and theories to which they give rise, which provides a solid basis in terms of which reality may be understood.

The elimination of metalogical projections comprises a first step in reaching a comprehensive understanding of reality. We gain knowledge of reality by *subtracting away* that which interferes, distorts, and is deceptive. This constitutes a major and significant advance by delineating the horizons in terms of which reality, in principle, is possible and can meaningfully be investigated. The resulting “silhouette” supplies a basis for understanding reality by dispensing with what not only is unnecessary, but what is metalogically self-undermining. The limits of possible meaning are accordingly established.

In a second step, by correcting and revising our fundamental conceptual vocabulary by means of the method of de-projection, we substantively *add* to our knowledge of reality, replacing concepts and claims that involve delusional reference and meaning with those that accord with principles of frame-

work relativity.

The elimination of projective concepts and claims, together with the formulation of a de-projectively coherent, corrected, and revised vocabulary of concepts, provides a sound, meaningful reflective framework in terms of which a *fundamental, unified, and comprehensive theory of the world* is made possible. We come to see that reality is, on the most fundamental level of possible reference and meaning, essentially coordinative or correlational, necessarily functionally defined in terms of the frameworks of reference in terms of which reference to objects is possible. We also come to see that to wish for more than this is, in principle, without possible meaning and rationally must be relinquished.

In short and as we have seen in previous chapters, the method of de-projection serves a negative, eliminative function, and yet it also serves an explicitly positive, constructive purpose when projective concepts are de-projectively corrected and revised. As noted in connection with the numerous forms of projection listed in the previous section, the de-projective analyses of previous chapters of this study have frequently led to fundamental revisions of initially putatively meaningful concepts so that they not only are no longer metalogically self-undermining, but are understood to have new and often counter-intuitive meanings.

For readers willing to relinquish concepts and claims that have been shown to be projective, and who are then able to integrate de-projected conceptual revisions in their thinking, substantively more than a “silhouette” of reality results: An understanding of reality in framework-relative terms is gained that provides a significant amount of conceptual content which a “silhouette” alone is incapable of supplying.

Certainly no book is capable of bringing about an integration in a reader’s understanding; this is something for which the individual thinker is responsible. In the foregoing chapters, many of the interlocking elements necessary for such an integrated understanding of reality have been described in some detail. The rest must be left to the reader.

### **30.4 Intrinsic limitations of reference and identity**

The negative science of the metalogic of reference brings to light the parameters of constraint that define the system of reference that *is* a particular theory or position, and therefore makes clear what is possible from that standpoint, and hence what from that standpoint is possibly meaningful. The main consequence to which the metalogic of reference leads is the rejection of what which has no possibility of being meaningful. The results of the metalogic of

reference are *essentially limitative*. As we saw in connection with theoretical physics ({26–28}), it is just such rejection that has provided the basis for the epistemological advances and many of the chief insights of special and general relativity and of quantum theory. In physics, these advances and insights, although theoretically abstract, are nonetheless given material content by the ultimate aim of physics to apply its mathematically formulated understanding to physical phenomena. The metalogic of reference, in contrast, remains on a purely conceptual level, investing its interest and energy in conceptual analyses, detached from concern for its application to the pragmatical uses which individuals may make of its results. The latter is a task that must be delegated by the critique of impure reason to teachers, teachers who ideally are trained and skilled in encouraging and assisting their students to overcome the human dispositions of willful blindness, belief system recalcitrance, and the other cognitive and psychological blocks identified in previous chapters. Further discussion relating to the challenge facing those teachers and students who wish to take on this task will be found in Appendix II.

Earlier in {4.5}, the conceptual level of the metalogic of reference was described in Fitch's terms as "vertical" or "non-ordinal." We recall that theories developed, for example, in empirical science have a low and finite ordinal level: Such theories have a concrete empirical focus, and are not themselves theories about theories. They are ordinal theories; they have a "horizontal" focus. In contrast, some theories are theories about all theories; they have no ordinal level; instead, we have characterized such metatheories as theories articulated on the level of maximum theoretical generality. This is a level of theory from the standpoint of which theories of all subordinate levels may be investigated. The level of maximum theoretical generality is maximal because there is no higher level from the standpoint of which it is possible to study theories that cannot be studied from the level of maximum theoretical generality.

The metalogic of reference is a self-referential theory, that is, it is itself included in its own proper subject matter. An ordinal level cannot be assigned to a self-referential theory, for the meaning we have given to the term 'ordinal' is inapplicable to such a theory. The level of theoretical abstraction appropriate to the metalogic of reference, then, is non-ordinal, vertical, and self-referential.

The metalogic of reference is, moreover, *maximally fundamental* in the sense described in {29.1}: Its focus upon the preconditions of identifying reference is conceptually the most fundamental possible since there is no more fundamental level of study that cannot be studied from the maximally fundamental level of identifying reference; for if there were to be such a pur-

portedly more fundamental level, it would presuppose the capacity in principle to refer identifyingly to its subject-matter.

Once the level of maximum theoretical generality and the maximally fundamental nature of the metalogic of reference are understood explicitly, clearly, and unequivocally, we recognize why there has been an inevitable progression in this study from the traditional, relational, extrinsic notion of reference, to an intrinsic, non-relational understanding of reference ({5 and 25}). In contrast to the traditional, commonsensical, relational, subject-object, agency-based notion of reference, which comprises an *extrinsic* model—permeated, as we have seen in Part III of this study, with projective notions—the metalogic of reference leads to a recognition of the framework-relative *intrinsic* functional interdependency of objects of reference with respect to the referential fields in terms of which they may be identified. The traditional extrinsic separation of object of reference from frame of reference gives way to the “absorption” of the constitutive structure of a given reference frame in the identity of objects to which that reference frame permits reference. In this sense, we have come to understand ({9.2, 10.2, 14.2}) that any object of reference has “ingredient” or “embedded” within it, as an integral constituent of its identity, the constitutive structure of the reference frame in terms of which it is identifiable. It is in this sense that we recognize that referential fields form interrelated systemic totalities so that, on the most fundamental level, a de-projective analysis comprises an affirmation of the system of interrelation that makes reference from the standpoint of that system possible. If the reader may not yet have come to an intuitive grasp of the preceding statement, or if it may continue to seem excessively abstract, he or she may be helpfully reminded of Einstein’s metaphorical “mollusk” ({26.8}) which in general relativity expresses much the same concept.

Understood in this manner, the identity of objects of reference and their very identifiability are functionally constrained by the frames of reference that permit reference to them. We have expressed this limitative result by means of the concept of horizons of possibility and meaning. The form of argument to establish this result has been self-referential, focusing on the preconditions of possible reference and meaning. We have come to recognize referential consistency as a necessary, intrinsically determined general criterion of meaning. We have affirmed the logical priority of the bonded pair, reference-and-meaning ({12.5, 13, 14.2}).

These results, taken together, have given new meaning to framework relativity: We’ve come to see that framework relativity involves both the self-enclosure of the referential field associated with a system of reference, and the embedment in the identity of any object of reference of its constitutive struc-

ture, namely, the preconditions necessary for its identifiability. The epistemologically neutral, de-projected understanding of reference we have come to ({15.2}) complies with the preconditions of reference that define any given framework of reference, and introduces no supplementary content or assumptions. We've seen that, as a heuristic method, de-projection makes it possible to describe, diagnose, and eliminate projections in a manner whose logical structure may be characterized as tautologous. We recall that the method is empty of content in the tautologous sense that no content is introduced that is not already ingredient in the system of reference under analysis. In making explicit the metalogical constraints that a given frame of reference presupposes, reflective analysis reaches a point where we are strongly, i.e., self-validatingly, compelled to see that the acceptance of a given theory or position metalogically entails the structural/systemic presuppositions of that theory or position, and with those presuppositions, a horizon of possibility and meaning is inevitably associated. In these terms, we have found that a dynamic, systems-based understanding of the modal concept of referential fields is a consequence of what meaningfully may be said to be. And, as a corollary to this, we have seen that to attempt to claim more than this is to fall victim to projective delusion.

### **30.5 *The Critique of Impure Reason* and conceptual revolution**

The prolonged failure of traditional means ... does not prove those means are useless. It does strongly suggest their inadequacy. For, as knowledge of the creative process drives us to conclude, although a problem which stubbornly resists solution by traditional means may perhaps be insoluble, the probability is rather that those means are themselves inadequate: the concepts, attitudes, and procedures employed are probably at fault and in need of being transcended in a fresh approach. The only reasonable step, at this point, then, is to act upon the supposition that our problems ... may be soluble only creatively—that is, by a profound and thorough alteration of our inner life and of the outer forms in which life finds expression and support....

[T]he fact is that there is a great deal of stability, so much that often it interferes with life. It may be that the threat of dissolution is so great that men have developed their conservatism as a necessary guard against the dispersal of the order they live by. Whatever the cause, the tendency to distrust the

widest and freest ranging of the mind is so strong that the changes necessary ... could not be attained without the efforts of the more daring and ingenious of mankind.

– Brewster Ghiselin (1985/1952, pp. 3, 13)

Intellectual history has been punctuated by occasional, successive revolutions in human thought. I take the history of conceptual revolutions in cosmology as one of the most important and significant examples of these revolutions, for cosmology serves to situate human life and human effort in the broadest context of space and time. In cosmology, a succession of great shifts in human thought has taken place, from the earliest conception of the Earth as flat to the Ptolemaic notion of the Earth as center of the universe with all else revolving around it. Some fourteen centuries later, there came the first true scientific revolution in cosmology, the Copernican-Galilean revolution, shocking humanity by demoting the Earth to one among other worlds circling in a planetary system around the sun, the heliocentric model. Several centuries later there followed a second conceptual revolution in the form of the space-time, field-based relativistic model developed by Einstein and then extended by others. These two giant steps required radical changes in thinking, each permanently affecting the course of later efforts. Each revolutionary step was initially opposed by the conservative anxieties of traditionalists, who did their best to fetter and block those who would break new ground.

In comparison, in the history of philosophy we have witnessed a succession of fashionable models, some more transient than others. Without a clear progressive, evolutionary history, it is much less clear that philosophy has experienced conceptual revolutions that can be compared with those in cosmology. However, if we wish, we may select certain of the major contrasting philosophical views and place them by analogy and in very rough parallel with the evolutionary steps of cosmology. If we do this, we may picture these steps: from the naive realism of the primitive savage whose world we may imagine consisted almost exclusively of physical events and conditions of nature, corresponding to the Earth initially seen as flat, and then over millennia broadening so that the experienced world of physical phenomena was thought to revolve around a central “I” of consciousness, a Ptolemaic analog. This innocent philosophically childlike view was then bifurcated by Descartes’ *cogito*, which inserted a wedge between subjective awareness and the rest of the world. Descartes’ separation of mind and matter brought with it a radical shift in philosophical thought, unsettling and to an extent shocking the discipline. Idealism received inspiration from the Cartesian division of mind



from physical reality and sought to contain the physical world within consciousness, while realism took its opposing stand. With the passage of time, a multiplicity of varied hybrid offspring has come into being, the gamut of the numerous philosophical positions that we find today.

Although the history of philosophy has witnessed a succession of many views, perhaps only Descartes' qualifies as a true revolution in thinking which has, in a persisting way, influenced a significant portion of the course of later philosophy. It is, I propose, time for a second revolution in philosophical thought, one that bears fundamental and non-accidental similarities to the second, the relativistic, revolution in physics.

The metalogic of reference is unmistakably revolutionary in reaching limitative conclusions that urge the renunciation of and radical departure from such traditional philosophical notions as "external reality," "subjectivity," "agency," "the self," "causality," "free will," "determinism," and other traditional philosophical conceptions, a renunciation extending even to the commonsensical subject-object, activity-based notion of referring. The metalogic of reference seeks to clear the table of a large vocabulary of conventional, familiar concepts which have been believed to be meaningful, in a manner insufficiently critical, by the great majority of philosophers. These are concepts that have become so habitual as now to be second-nature, but they are concepts, nonetheless, which, as we have seen, undermine their own possibilities of reference and meaning.

The metalogic of reference evidently requires a new way of thinking. A new way of thinking has seldom in human history been welcomed. Russell once commented, "[i]n spite ... of the ... possibility of progress in philosophy, the first effect, as in the case of physics, is to diminish very greatly the extent of *what is thought to be known*" (Russell, 1972/1914, p. 243, italics added). The eliminative objective of the negative science developed in these pages does indeed "diminish very greatly the extent of what is thought to be known." It does this through its critique of a large group of putatively meaningful concepts which, one after another, have been shown to be self-undermining on the level of possibility of reference and meaning.

This study has made a bold group of claims that can be expected to touch the conservative and defensive instincts of any philosopher or other reader who is made uncomfortable by such a challenge to his or her routine and habitually accepted conceptual vocabulary, a vocabulary which he or she *thought was known* to be meaningful.

To be willing fully to engage in such a critique of impure reason requires, as I commented early in this study, not only patience and fortitude on the reader's part, but a willingness genuinely to listen to and understand another

philosopher's position, plus an honest desire to reach an agreed-upon unitary set of criteria and methodology that can be called upon to judge the validity and acceptability of any philosophical position. What is more, such a critique requires the possession of an intellectual-psychological constitution that enables, encourages, and in fact compels the reader to cross the bridge of rationality and commit to the conclusions which rationality demands ( $\{1.2\}$ ). More than these dispositions, however, it is necessary—whenever a reader is called upon to think in a new way—to possess a readiness to “think outside the box” of accepted and customary concepts. (For more relating to this disposition and necessary associated skills, see Appendix II.) As Nietzsche expressed this with his usual dramatic flare: “He must have an inclination, born of strength, for questions that no one has the courage for; the courage for the *forbidden*” (Nietzsche, 1918/1895, p. 37).

### 30.6 A possible future of philosophy

The one and only condition, I believe, which is necessary in order to secure for philosophy in the near future an achievement surpassing all that has hitherto been accomplished by philosophers, is the creation of a school of men with scientific training and philosophical interests, unhampered by the traditions of the past, and not misled by the literary methods of those who copy the ancients in all except their merits.

– Bertrand Russell (1972/1914, p. 246)

More than one hundred years have passed since Russell expressed this hopeful belief, and that hope evidently remains unfulfilled. We are as far as ever from having produced a school of men and women with scientific training who are able to be free from the ingrained accustomed ideas of the past and the routinized ways of thinking that employ those ideas. Instead, the reaction must be one of dissatisfaction on the part of anyone who is dominated by the need in the discipline of philosophy for real substance, for demonstrable content, for reliable results upon which future efforts can incrementally be based. This is no less true even for less demanding students and readers of philosophy who would gain trustworthy insights applicable to their outlook and understanding of reality. The justifiable reaction of such a person to the ideas of philosophers as they continue to write and publish their papers and books was articulately expressed by Tolstoy's down-to-earth sincere character Levin in *Anna Karenina*:

Their ideas seemed to him fruitful when he was reading or was himself seeking arguments to refute other theories, especially those of the materialists; but as soon as he began to read or sought for himself a solution of problems, the same thing always happened. As long as he followed the fixed definition of obscure words such as *spirit*, *will*, *freedom*, *essence*, purposely letting himself go into the snare of words the philosophers set for him, he seemed to comprehend something. But he had only to forget the artificial train of reasoning, and to turn from life itself to what had satisfied him while thinking in accordance with the fixed definitions, and all this artificial edifice fell to pieces at once like a house of cards, and it became clear that the edifice had been built up out of those transposed words, apart from anything in life more important than reason. (Tolstoy, 1919, pp. 990-991)

Philosopher Walter Kaufman similarly wrote: “[W]hat is the use of studying philosophy if all that it does for you is enable you to talk with some plausibility about some abstruse questions of logic, etc., and if it does not improve your thinking about the important questions of everyday life...” (Kaufman, 1963, pp. 36-37).

In Part I of this study we discussed some of the ramifications of such discouraging observations concerning the discipline of philosophy. My purpose there as well as here is not to devalue the profession, only to appraise and perhaps contribute to its efforts to establish reliable and enduring results upon which future similarly conclusive findings can be built. Its success to date has been less than encouraging. Is there a possible future in which this might be remedied?

We can, I think, still be optimistic that a scientific approach may come to be regarded as the preferred arbiter of philosophical research. A horn of plenty has been filled, in particular, by the fruits of the labor of logicians, whose systematic proof-oriented approach may offer the most promising avenue to definitive conclusions. Unfortunately, much of this work remains technically inbred, evidencing little general interest in developing a comprehensive approach that can unify the fragmented multiverse of individual competing systems and positions. There is a need to which few philosophers respond, to develop a single, unitary, all-inclusive theoretical framework governed by the norms of science in terms of which ongoing and past work can be assessed.

The approach presented in this book seeks to contribute to that objective.

To the extent that traditional philosophical thought, as well as much of the thought of scientists, mathematicians, humanists, and the general public, is encumbered by delusion-inducing, self-undermining projective notions, it is indispensable to begin with a thorough conceptual “house cleaning.” As so many revolutionaries have claimed, and as philosopher Gaston Bachelard urged, “before building, one must destroy” (Bachelard, 1959/1949, p. 572). To be willing to proceed without such an initial and systematic discipline of error elimination is to continue to serve as a carrier of errors, allowing them to contaminate and hold back future work. No physician can hope to heal if his medical instruments are allowed to remain undisinfected. There is an unavoidable need, if philosophy is to make actual progress, for the negative science of a critique of impure reason. Without it we shall be left with a discipline characterized by fashion-driven, indemonstrable contentions of the multitude of often competing, branching sub-specialties that continue to proliferate in philosophy. As Albert Schweitzer observed:

I had all along felt it to be psychically a danger that in the so-called humanities with which I had been concerned hitherto, there is no truth which affirms itself as self-evident, but that a mere opinion can, by the way in which it deals with the subject matter, obtain recognition as true. The search for truth in the domains of history and philosophy is carried on in constantly repeated endless duels between the sense of reality of the one and the inventive imaginative power of the other. The argument from facts is never able to obtain a definite victory over the skillfully produced opinion. How often does what is reckoned as progress consist in a skillfully argued opinion putting real insight out of action for a long time!

To have to watch this drama going on and on, and deal in such different ways with men who had lost all feeling for reality I had found not a little depressing. (Schweitzer, 1949/1933, p. 104)

### **30.7 Intellectual humility: Submission of philosophers to the norms of science**

The set of cognitive and psychological dispositions that characterize many philosophers, as described in Part I, places in relief the fact that, as a professional group, philosophers highly value independence of mind and the abilities to think clearly, analytically, and critically. They hold fast to the intel-

lectual freedom that their discipline underwrites, which sanctions them to build and cultivate their own independent philosophical positions. These values do, I think, stand for goals that in some ways are exemplary and important. Unfortunately, as I have tried to emphasize, their exemplification in the actual work and interpersonal communications of philosophers is far from ideal. In the reality of contacts between philosophers, whether in person at professional conferences or in philosophically focused exchanges or in reactions to one another's printed work, their prevailing personality is self-absorbed, defensive, and intent upon supporting their own individual preferred interests and positions, and securing these from attack by other philosophers. This prevailing personality is very plainly proud and territorially self-interested and self-invested.

In stark contrast, students of science and its professional practitioners tacitly agree upon and accept the defining norms of their discipline. Among scientists, it is rare for these norms to be made the explicit subject of discussion, and even more seldom are those norms questioned. The standards of justification and the standards of acceptable scientific practice are learned and internalized as science students progress through their studies, and the norms that define science become for them a routine, conventionally accepted and endorsed backbone of all scientific endeavor.

To urge that philosophers agree to submit to such norms would require persuasion, argumentation, dialogue, contention, and debate, which, taken together, may predictably be expected to result in the formation of some groups whose members respond, to some extent, favorably, and other groups whose members may well resent and reject the suggestion outright, and may feel affronted by any perceived pressure placed upon them to "mimic the ways of science." An impasse would surely result.

It is not my purpose to make a case for a single approach leading to an exclusionary conception of philosophy as a genuine science. There is space enough in the world of the mind for a diversity of approaches. Unfortunately, very little space has been occupied by philosophers who wish for more in the way of definitive and enduring results by their discipline than it has so far been able to produce.

To enlist among a meaningfully large group of philosophers the interest and desire to contribute to philosophy as a genuine science requires a shared willingness to submit to the norms that govern scientific thought, research, and publication. Given the intellectually proud, independent, position-taking propensities of, in my experience, the majority of philosophers, such a willingness to submit to the standards that define scientific practice requires a strong measure of *intellectual humility*. To be willing to be governed in one's

thought and work by a unitary set of criteria of justification does demand a kind of humility, of modesty, which is certainly intellectual in nature, but also personal. It is a form of humility that accepts that one's efforts and resulting work are appropriately and validly to be judged using standards universally shared by others in the scientific community.

This kind of humility never becomes a conscious issue for those who enter and pursue scientific careers. It does, however, become an issue for any philosopher who develops aspirations that his or her discipline might come to include a genuinely scientific orientation, with a shared and conceptually rigorous methodology and standards of strict justification.

Philosophers of the past who have in their various ways urged what they have considered to be a truly scientific approach to philosophy—for example, Bertrand Russell, Hans Reichenbach, Edmund Husserl, Rudolf Carnap, Moritz Schlick, and others—have, as I see it, ignored the *preferred, self-chosen* limitations of their position-taking fellow philosophers, and, as a result, have to a great extent, so to speak, “talked past them.” The issue of the special variety of intellectual humility that is a prerequisite for philosophers who would consider subjecting themselves to the norms of science is never mentioned.

I have mentioned it here, and have wished to emphasize the central role of such intellectual humility before proceeding to the next section which is written in the spirit of that form of humility.

### **30.8 The logical standing of the method and results of the metalogic of reference**

By ‘*logical standing*’ I mean the relative rank or position on a scale of degrees of certainty of the principal conclusions we have reached in this study. We may think of logical standing as a rank or position that can be placed on a spectrum extending from the arbitrariness of unfounded opinion and associated uncertainty, to conclusions reached by means of demonstrable proof. We may think of grades of certainty, which culminate in valid derivations within a consistent formal system or in the certainty of a tautological equivalence.

The logical standing of the method of de-projection and of the de-projective solutions reached in this study is *definitive*, in two special and essentially interrelated senses; I will call each a ‘dimension’ of logical standing: In the first sense, the method of de-projection and the de-projective results reached by means of its application are inescapably compelling: To reject either the method itself or the de-projective results reached is to fall victim to metalogical projection, undermining—in principle—the very *possibility* of such a re-

jection. The method and the de-projective results reached are self-validating in the special sense that has been developed in this study: They cannot *not* be accepted without undermining the very possibility of reference and meaning from the standpoint of the frame(s) of reference that is/are at issue and must be presupposed. The logical structure of both the method of de-projection and of the de-projective results reached through its application is ultimately tautologous, involving a re-affirmation of and compliance with the constitutive structure of any given coherent frame of reference. The chapters making up Part II of this book claim to demonstrate this.

There is also a second sense in which the logical standing of the method of de-projection and the results reached is definitive. This is the sense in which, if a philosopher has a mind equipped with what I've called 'epistemological intelligence', he or she will be entirely convinced of the validity of the results reached once the path to those results has been studied carefully. For our purposes here, it will not be necessary to define the concept of epistemological intelligence in any detail (for readers interested in this concept, see Appendix II). Here, let us simply associate with this concept an imagined group of largely cognitive skills that are indispensable to valid epistemological analysis. The logical standing of the main results of this study are, in this second sense, relative to the framework established by this set of skills. In other words, without those skills, a reader will very likely not be strongly compelled to accept the results of the analyses found in these pages. This is a variety of cognitive framework relativity, the role of which it is important not to ignore.

This second dimension of logical standing, taken in conjunction with the observations made in Part I of this book concerning the psychological-cognitive profile of many philosophers, constitutes what is essentially a caveat: For without the second dimension, it is unlikely that the full conceptually compelling force of the first dimension of logical standing will be appreciated or accepted by some readers. The two dimensions that define the logical standing of the results reached in this work are in this way essentially interrelated.

To offset the suspicion that this caveat camouflages something rather mysterious, it does not. It is a caveat that usually goes unnoticed when we speak of the results of any demonstration. Russell made it quite explicit in speaking of then-recent advances in mathematical logic:

[T]he new logic ... has ... introduced the same kind of advance into philosophy as Galileo introduced into physics, making it possible at last to see what kinds of problems may

be capable of solution, and what kinds must be abandoned as beyond human powers. And where a solution appears possible, the new logic provides a method which enables us to obtain results that do not merely embody personal idiosyncrasies, but must command the assent of *all who are competent to form an opinion*. (Russell, 1972/1914, pp. 68-69, italics added)

The second dimension of logical standing that I have distinguished above is expressed in abbreviated form in the above italicized phrase.

The results this study has reached are claimed to have the logical standing that I have described. Their logical standing includes a third and final “dimension,” which is this: Where it may perhaps be found that any of these results require correction or modification, the further claim made by the self-enclosure of the metalogic of reference is that such corrections or modifications will, of necessity, come to light by presupposing the very metatheoretical means employed and affirmed in this study. In other words, the possibility, in principle, of detecting the need for such corrections or modifications, as well as the possibility of making such corrections and modifications, is insured by the reflexive, vertical, non-ordinal framework of the metalogic of reference, a framework developed on the level of maximum theoretical generality.

In the words of mathematician Paul Halmos: “This settles all our problems and fulfills all our promises” (Halmos, 1972/1951, p. 109).





## SUPPLEMENT

### **The Formal Structure of the Metalogic of Reference**

Gödel anticipated the development of “philosophy as an exact theory” which “should do to metaphysics as much as Newton did to physics.... [I]t is perfectly possible that the development of such a philosophical theory will take place within the next hundred years or even sooner.”

– Hao Wang (1987, p. 192)

**I**n this Supplement, the question of the extent to which the metalogic of reference can be formalized is discussed, followed by a formalized expression of certain of the main concepts and properties that define its logical structure, simplifying and to an extent amplifying {11.4}. This Supplement provides a conceptually abridged and compact formulation of a group of the basic principles of the metalogic of reference. It is not intended to serve as interchangeable with or as an adequate substitute for the contents of the main text, but as a potentially informative, supplemental means of understanding the nature of the metalogic of reference.

...

The possible representation in a formalized language of the formal structure of the metalogic of reference presents a set of conceptual challenges. As a vertical, non-ordinal, self-referential theory developed on the level of maximum theoretical generality, the metalogic of reference possesses a formal structure that does not easily lend itself to the standard tools of mathematical logic and the theory of formal systems. It is important that we be aware both of the advantages of those tools as well as of certain of their chief limitations in order that we may approach the question of formalization with an informed sense of what we can and cannot expect.

### §1. The advantages and the shortcomings of formalization

It is often more easy [sic] to express correctly a formal system than the philosophy behind it. Sometimes this is because the formal system does not express the philosophy adequately. Then we find the system objectionable because of the discrepancy. At other times a system may happen to be better than the philosophy behind it. There are also interesting systems behind which there is little articulate philosophy.

– Hao Wang (1953, p. 133)

Generally speaking, the ways in which mathematicians and philosophers go about thinking are very different, and this is particularly true in connection with formalization. Those mathematicians whose chief interest is formalization, and those philosophers whose concern is to formulate abstract principles of understanding, proceed in their thinking in very nearly opposite directions: A mathematician who wishes to develop a formalized system generally attempts to lay the mathematical groundwork by means of elementary propositions, or axioms, that are subsequently to be taken for granted; these are accompanied by a set of rules of formation of expressions along with a set of rules of derivation which, when applied to the initial elementary propositions, make the derivation or deduction possible of many other propositions of increasing complexity which follow as logical consequences. The direction of such a mathematician's thinking proceeds, then, from elementary to increasingly complex sophistication.

In contrast, a philosopher who is intent upon identifying highly broad or universal generalizations about a subject-matter often begins with information acquired and refined from the complexities and the manifold variety of ordinary experience, and attempts to take note of patterns of commonality or difference that characterize such experience, and from the patterns he or she recognizes proceeds then to formulate principles which plausibly or by means of argument lead to a clearer, more insightful, or more penetrating understanding of that subject-matter than was originally possible. The direction of such a philosopher's thinking proceeds, then, from the original unanalyzed, interwoven, detail-laden multiplicity of experience to an increasingly simplified, abstract expression. This direction of philosophical thinking is virtually the reverse of the mathematician's.

The question whether or not it can be advantageous for philosophers to formalize their arguments and positions has no single satisfactory answer applicable to all. A satisfactory answer will depend upon the nature of the

individual philosophical position and the varieties of argumentation employed by the philosopher in an effort to establish the position's validity or persuasiveness. In the case of the metalogic of reference, to which I will soon restrict this discussion, the question of formalization cannot, as we shall see, be answered easily or as yet with finality.

### §1.1. The advantages

On the side of the advantages of formalizing a philosophical position or argument a good deal can be said that is positive. Very much as an anatomist gains insight into the workings of the human body by having an assembled skeleton available for study, so does the successful formalization of a philosophical position or argument exhibit its underlying structure, dissected out, so to speak, from the original, often heavily laden, complexifying mass of extraneous and obscuring detail, of multi-layered conceptual and concrete meaning, all of which can distract from an unencumbered, clear assessment.

Philosophical thought, as it generally is to be encountered in its published literature, in the oral presentations of philosophers of their work, and in exchanges among individual philosophers, tends to be a messy affair. It can be hard to "dissect out" what is actually vital to an argument or the position a philosopher propounds. When it is possible to give an adequate formalization of the argument or position, much needless wild goose chasing may be avoided. Philosophical thinking and its expression seldom are rigorously controlled undertakings; as Max Black (1959/1933, p. 142) once commented, "thought is always ahead of adequate symbolization." —Perhaps not always, but usually this is true.

By formalizing a philosopher's argument, or his or her entire system (if such there be), the reader's task can be greatly simplified, making it possible to identify the philosopher's most basic premises and the rules of reasoning upon which the rest is supposed logically to follow. It may then be possible to prove that the application of the proposed set of rules of reasoning to the basic premises will not lead to contradiction, or to prove that the set of basic premises is complete, permitting all desirable true propositions to be derived from them, or to show that each of the premises is logically independent and cannot be dispensed with, or to show that the set of premises is minimal in number. Furthermore, formalization can sometimes facilitate a philosopher's reasoning when dealing with highly abstract concepts, which otherwise might be more difficult or complex were only natural language to be relied upon. At times, formalization itself may lead one to conceive of possibilities which informal reasoning can miss. Formalization has unquestionably great value in allowing

one to home in on an argument's or a position's potential weaknesses, and in directing attention to its potential strengths. A formalization can be highly useful as a means of understanding the relationships claimed to hold among the principal concepts in terms of which an argument or position is developed. The clear advantages of formalization led logician Arthur Prior to state categorically: "it is important to 'formalise' as much as we can, i.e. to state truths about things in a rigorous language with a known and explicit structure" (Prior, 1996, p. 45).

All of the advantages I have mentioned are well-known benefits of formalization. With the development of the multiplicity of tools of modern mathematical logic and the rapid evolution of computer science and artificial intelligence, formalization has acquired an almost uncritically accepted and endorsed *cachet* of its own, at times unjustifiably and deceptively appearing to elevate the mundane to sublime heights. Therefore, to balance this sometimes blind allegiance to the spirit of formalization, let us also look at a group of its major shortcomings.

### §1.2. The shortcomings

One of the most perceptive and candid discussions of the question of formalization as it relates to philosophy is found in mathematician Hao Wang's short paper (Wang, 1955). His remarks cut to the core—or, as we shall see, at least *part* of the core—of the issue:

The application of mathematical logic to the treatment of philosophical problems may ... be viewed as an attempt to formalize. Such applications often give the impression that a formidable technical book expresses in tiresome exactitude more or less commonplace ideas which would be conveyed more easily and more directly in a few sentences of plain language....

Perhaps we can compare many of the attempts to formalize with the use of an airplane to visit a friend living in the same town. Unless you simply love the airplane ride and want to use the visit as an excuse for having a good time in the air, the procedure would be quite pointless and extremely inconvenient. Or we may compare the matter with constructing or using a huge computer solely to calculate the result of multiplying seven by eleven. (Wang, 1955, p. 233)

This is, self-evidently, one of the shortcomings of formalization: Formalizing a philosophical position or argument is often simply unnecessary and can complicate with technical sophistication rather than simplifying that which it seeks to formalize. George N. Schlesinger extended Wang's remarks: "Some have gone even further to charge that trivial issues are often made to look more impressive when blown up by formal contrivances to a size much in excess of their real importance" (Schlesinger, 1985, p. xiii).

A second shortcoming of formalization is the unfortunate result it may have both for the formalizer and for his or her audience. By formalizing an argument or a position it can happen that, by the very fact that a formalization has been accomplished, the possibility of reaching a genuine understanding of a problem can be displaced and then replaced by the mistaken belief that one has reached an understanding purely *because* that formalization has been developed. This is to make the error in thinking that, having formalized an argument or position, one has therefore understood it. One of the first to have recognized this mistake was mathematician Allan Calder when he wrote:

[C]onstructivists try to avoid formalizing their theories, because once they are formalized it is possible to generate theorems *without the need to understand what is going on*. You do not have to understand a proof to know that it is formally correct....

Most of mathematics that I do is highly constructive, but I do feel strongly that formalism must be kept in a proper perspective and not used *as a substitute for understanding*. In my opinion formalism is the opium of the thinking classes. (Calder, 1980, p. 6, italics added)

A third shortcoming of formalization, closely related to the second, is that formalizing an argument or position may fail to capture the meaning of what is at issue—sometimes because the formal language chosen is insufficiently expressive, and sometimes perhaps (one may conjecture) because no formal language may, in principle, be able adequately to represent the meanings that are involved in the subject under consideration. We shall discuss this possible and important shortcoming in the next section.

A fourth potential shortcoming is unmistakably damaging and can effectively undercut a philosopher's motivation to formalize. It concerns the possibility that ordinary, everyday reasoning, and also the more complex, multi-leveled informal reasoning indispensable to much philosophical thought, may not be formalizable in any deductive sense. We shall return to this potential

shortcoming later.

Fifth, there is yet another and very obvious disadvantage that can result from formalization: The prerequisite technical sophistication on the reader's part to understand it can defeat a philosopher's efforts to communicate to more than a specialized audience.

## §2. Internal limitations of formalization

The third potential shortcoming identified above can be a manifestation of certain of the *internal limitations* of formalization. This shortcoming implicitly points to the significant question whether formalization may—perhaps in principle—fail satisfactorily to translate into the language of a formalized system a particular subject under philosophical consideration. Given the distinctive and non-standard defining properties of the metalogic of reference, it is important that we take clearly into account limitative results that may, in principle, stand in the way of its formalization.

A number of the limitative results that apply to formalized systems have now become well-known to many philosophers, as they have become second-nature to mathematical logicians. In this brief Supplement, I cannot of course describe the series of theorems of formal limitation proved by the numerous contributors to this area of study, which have included Gödel, Rosser, Tarski, Wang, Post, Kleene, Curry, Löwenheim, Skolem, Henkin, and others. For readers who wish to deepen their knowledge in this area, in my judgment the finest, most comprehensive, and most insightful work was written by Belgian mathematical logician and philosopher Jean Ladrière (1921-2007), *Les limitations internes des formalismes* (Ladrière, 1992/1957).

In what follows, I limit my discussion to a small group of philosophical ramifications of the theorems of formal limitation insofar as these potentially apply to the metalogic of reference.<sup>360</sup> This means that it will be possible here only to highlight in abbreviated form a few of the limitative results which, in different ways, constrain the extent to which the deductive languages of formal systems can be employed to express the formal structure of a highly general metatheory of the kind developed in this study. For this discussion I am indebted to Ladrière's opus, which has never been translated into English and as a result is largely unknown to Anglo-American philosophers and

---

<sup>360</sup> For reasons of space and of direct relevance to the metalogic of reference, I exclude in this discussion non-Gödelian, non-classical approaches to formal systems, which renounce certain of the conditions of deductive effectiveness, or at the expense of admitting principles such as indefinite extensibility and/or essential indeterminacy.

logicians.<sup>361</sup> I will take the opportunity to quote several passages, translated into English, from Ladrière's book which I think throw light on how the theorems of formal limitation relate to the formal structure of the metalogic of reference, and hence how its structure is to be understood from the point of view of formalization.

Of the various limitative results, best known among philosophers is probably Gödel's theorem in its different forms,<sup>362</sup> its generalization, and theorems related to it. Let us take those results as an example: They show that any formal system of a certain minimal degree of complexity contains propositions which are true, but yet which are undecidable in that system. Gödel formulated such a proposition using the language of a given formalized system (capable of containing arithmetic, e.g., the system of Russell and Whitehead's *Principia*); this proposition, in effect, affirms of itself that it is not derivable in that system. And yet this proposition is true—neither is it derivable nor is it refutable given the proof resources of the system. But this is a fact accessible only by employing a mathematician's metatheoretical reasoning: In other words, a recognition of its truth relies upon the reflective capacity to consider a given formal system from a metatheoretical point of view. The preconditions of reference that must be satisfied in order for it to be possible to show that the proposition is true cannot be satisfied within the formal system in whose language the proposition is expressed: Recourse to a metaframework is necessary.

The theorems of limitation are metatheoretical theories. As such, they imply that it is possible to establish a distinction

---

<sup>361</sup> Ladrière's is a massive 700-page impressive and penetrating study. The fact that it had not been translated into English perplexed me, and I therefore discussed this state-of-affairs with Ladrière in 1971. I felt—as I continue to feel today—that the unfortunate denial of the work to English-reading logicians and philosophers should be corrected, and I therefore offered to translate the book into English. Ladrière, however, expressed his feeling that English-reading mathematical logicians were already in possession of the majority of publications in this field, and so he declined my offer, preferring to keep the book as a French language contribution to mathematical logic. (However, in 1969 he consented to the book's translation into Spanish, I believe because he recognized that the literature of mathematical logic available in that language was not extensive.)

Readers who wish for informative discussions in English of the theorems of formal limitation, may want to consider, for example, Smith (2007), Hájek and Pudlák (1993), and the more advanced Murawski (1999).

<sup>362</sup> His two incompleteness theorems are: First: Any consistent formal system of a level of complexity of elementary arithmetic is incomplete—that is, there are statements that can be formulated in the language of that system which can neither be proved nor refuted in it. And second: For any consistent formal system of a level of complexity of elementary arithmetic, its consistency cannot be proved in that system itself.

between theory and metatheory, and to give, in each concrete case, a precise meaning to this distinction.... [T]he metatheory of a system  $LF_1$  can be formalized in a system  $LF_2$ , and then in turn become an object of a metatheoretical study which could then eventually be formalized in a system  $LF_3$ . A metatheory can itself therefore be considered as a theory from the standpoint of another metatheory. But what is essential is that it is possible to undertake, in each case, a metatheoretical study of the envisioned system and that these metatheoretical considerations relating to the system are situated outside the system itself (then they can, in certain cases, be formalized—at least in part—in the system). (Ladrière, 1992/1957, pp. 415-416)<sup>363</sup>

This reliance upon metatheoretical—alternatively often called ‘metamathematical’—reasoning is a common and fundamental characteristic not only of Gödel’s theorem, but of the theorems of formal limitation developed by other mathematicians and logicians. After more than 400 pages of detailed description and analysis of the extended family of theorems of formal limitation, Ladrière summed up this essential reliance in the following words: “The theorems of limitation correspond to metatheoretical properties” (p. 434). The indispensable role of metatheoretical reasoning will be central to our discussion here.

### **§2.1. The impossibility of comprehensive or total formalization**

Comprehensive or total formalization is a mathematical ideal, and it *can be* only an ideal, one which, as it turns out, is in principle not realizable. If one were to try to formulate a single truth to which the various theorems of formal limitation lead, it is this. The ideal of comprehensive formalization is a limiting idea in mathematics that would be reached were it possible to develop an ideal formalized system that could be described in the following terms:

The ambition of the task of comprehensive formalization is the development of a system which contains in itself the genesis of its own meaning and which can therefore be considered autonomous.

---

<sup>363</sup> This and later translations from this work are by the present author. When there is no ambiguity, subsequent references to Ladrière’s book are given only to its page numbers.



Such a system would encompass the entire field accessible to [mathematical] intuition but at the same time it would be completely detached from any reference to intuition. Summarizing the acquisition of the totality of mathematical thought, it would contain in itself the key to all subsequent advances. Once constructed, it would in a way begin to proliferate by itself, progressively discovering all the virtualities inscribed in the field of the deductive. Uniformly applying established procedures once and for all, it would also make superfluous any true initiative of thought: it would offer an operational framework within which any problem could receive a mechanical solution. Finally, it would be capable of reflecting itself completely within itself: it would provide all the procedures necessary to formulate and solve the problems that could be posed about it; *it would for itself be its own metatheory*.

What the project of comprehensive formalization aims at is therefore the constitution of the total mathematical object. Because when the intuitive field is fully absorbed in the system, it can no longer be considered as an instrument. Having become coextensive with the reality which it was a question of reaching, it is no longer distinguished, ultimately, from this very reality. (pp. 408-409, italics added)

This highly abstract and idealized conception of comprehensive or total formalization bears certain unmistakable resemblances to the metalogic of reference. Like the metalogic of reference, the limiting idea of formalized mathematical comprehensiveness would establish a referential system whose level is that which we have termed ‘maximum theoretical generality’; that system of reference “would for itself be its own metatheory”: In the terms of the present study, it would have a wholly *intrinsic structure*, that of a *self-enclosed referential field*. The mathematical ideal would furthermore “be capable of reflecting itself completely within itself” and so it would be capable of complete reflexivity. These are all important commonalities shared by the metalogic of reference and the ideal of total formalization. And yet there is an important—and it may constitute a crucial—difference: We have not yet answered the question whether, or to what extent, the formal structure of the metalogic of reference may, in principle, be capable of formalization. The mathematical ideal is, of course, an ideal whose realization requires formalization. This requirement may be inapplicable to the metalogic of reference.

Before we attempt to respond to this issue as it relates to the metalogic of reference, let us reach the conclusion to be drawn from the theorems of formal limitation as they apply to the mathematical ideal of a comprehensive formalized system. The following passages express the conclusion Latrière himself reached. He refers to mathematical “intuition,” to the “intuitive field” of mathematical reasoning. Such intuition should be understood especially in the sense of metamathematical reasoning, the metatheoretical reasoning upon which the various theorems of formal limitation rely.

The total [formalized] system is not realizable as an adequate representation of the intuitive field, nor as a formal structure capable of reflecting itself entirely in itself, nor as a set of canonical procedures capable of providing an effective solution to any mathematical problem. (p. 411)

[W]e cannot exhaust in the present all the possibilities of reasoning which are, in principle, accessible to us. There does not exist a closed system which would be like the paradigm of all discourse, the supreme canon of reason. Whatever system we may consider, there are always forms of reasoning which remain foreign to it. (p. 412)

The possibilities of mathematical thought are not exhausted in the field of the decidable. There is therefore room for forms of reasoning which cannot be reduced to uniform and always effective procedures of the type which can be represented by means of machines. There is more in thought than can be enclosed within the exact limits of a calculus.

To say that there is no universal calculus is to say that we cannot reduce everything to purely formal rules of manipulation, that the domain of meaning is not to be equated with that of the effectively practicable and that, even in formalizations [*instaurations formelles*], contact with the sources of intuitive meaning remains essential. (p. 413)

[F]ormalism cannot adequately recover the content of intuition and, in this sense, the idea of a total formalization should be considered unrealizable. (p. 438)

The theorems of formal limitation have proved a group of metatheoretical properties which are inalienably tied to, and which necessarily constrain, the limits of classical mathematical formalization. They are “internal” limitations because they have been shown to be constraints that are entailed by the nature of such formalization itself. Our concern here is to consider whether these limitations also apply to the metalogic of reference, for, if they do, the formal structure of the metalogic of reference is not susceptible, in principle, to comprehensive formalization.

### **§2.2. The impossibility of the total reflexivity of formalized systems**

Related to the preceding section’s discussion is an important group of theorems of formal limitation that circumscribe the limits of reflexivity of formal systems. That group includes Rosser’s generalization of Gödel’s proof, Tarski’s theorem relating to the semantical concept of truth, theorems by Rosser and Wang, and related works.<sup>364</sup> Stated generally, these theorems demonstrate that all of the properties of a formalized system which can be recognized from a metatheoretical standpoint cannot, in principle, be expressed in that system; in other words, as affirmed by Ladrière, a formal system cannot undergo a total reflection of itself into its own system (p. 398).

The referential capacity of the metalogic of reference reflexively to take stock of its own properties and to express this self-referential metatheoretical self-awareness is, as we have seen in earlier chapters, essential to its purpose. As in the preceding section, we again encounter support for the contention that the structure of the metalogic of reference may not, in principle, be capable of comprehensive formalization.

### **§2.3. Formalization, temporality, recursion, and the metalogic of reference**

In earlier chapters, we found that the structure of identifying reference is temporal: Identifying reference to an object, of no matter what kind, is only possible when that object of reference is in some way fixed or defined within its

---

<sup>364</sup> For a detailed description and analysis of these theorems, see Ladrière (1992/1957, §127 enoncé XX, §128 enoncé XXI, §202 enoncé LXXXVI, §207 enoncé LXXXVIII, and §220). For further discussion of certain of the issues relating to the semantical conception of truth and the capacity of a formalized system to represent its own truth concept, see Bartlett (1970, Section 1.10, pp. 113-115 (English edition); Section 1.10, note 26, pp. 182-184 (French edition)).

presupposed framework of reference so that *re-identification* of that object is possible ( $\{10.1\}$ ), making it possible to recognize that one and the same object is in view from the standpoint of a plurality of referential contexts. For this to be possible, the re-identifiability of objects of reference requires that a reference frame permit *temporally successive*, reiterated reference in a manner so that retrospective second-order references are possible to past references. As we saw in  $\{21.1\}$ , this is an alternative way of expressing the metalogical fact that the possibility of reflection presupposes the means to retain temporally earlier states. In short, identifiability itself presupposes the possibility of reiterated reference, which in turn requires that retrospective, second-order, references be possible to past references. The temporal basis of reference relates, then, to its capacity for sequential or serial retention of earlier states.

The temporal nature of identifying reference, the basis for the successive reiterability of reference, has a fundamentally recursive character. For a function in mathematics to be defined recursively is in essence to provide a definition by mathematical induction. Typically, a recursive definition is given by means of a set of equations; one specifies the value of the function when its argument  $k = 0$ ; other equations specify the function's value when the argument  $k$  is incremented by 1. In this way, the value of the function for any given argument can be computed starting with its value when  $k = 0$ , and applying the other equations recursively—that is, successively—as many times as necessary.

The successor function is a simple example: The successor function,  $f$ , is recursively defined by the rule  $f(x) = x + 1$ : The function is defined by means of the rule which specifies the “next” element in the series. An element of a set of elements in a series is given, together with a rule which determines the next element on the basis of its predecessor.

Now, in one way or another, the theorems of formal limitation employ variations of the “diagonal procedure” first developed by Cantor. A diagonal procedure demonstrates the impossibility of a certain hypothesized correlation of elements by showing that such a correlation would permit the specification of an element which the proposed correlation could not cover. The demonstration exhibits an extra case which serves to refute the hypothesis of correlation.

The theorems of limitation show that recursive procedures of a formalized system of a certain minimal level of complexity cannot reach what metatheoretical reasoning itself is able to realize. Once a formal system possesses that minimal degree of complexity, metatheoretical reasoning can make recourse to Cantor-inspired diagonal reasoning to show an “extra case” which could

not be attained using the recursive procedures available in the given system. Ladrière states this succinctly:

One can consider that the notion of *decidability* is adequately represented by that of *recursion*. The theorems of limitation therefore teach us that recursive procedures cannot exhaust the field of the metatheoretical. And the precise reason is that a formal system of a certain degree of complexity can always make recourse to diagonal reasoning which, starting from an enumeration of integers of a certain type, shows that there is (at least) a set of that type which is not included in the enumeration. The phenomenon that we thus discover at the basis of all the theorems is that it is not possible to give an effective determination of all the sets of integers....

If formal systems present limitations, it is because they make recourse to diagonal reasoning, and it is therefore because their structure corresponds (in accordance with the modes of correspondence that are different in accordance with the meaning that one associates with the elements of the system) to that of the set of series of integers. This set possesses the remarkable property of not being exhaustible by enumeration, of being somehow indefinitely extensible, and of not being able however to lead us to an effective totalization of all its elements. (pp. 427, 432)

As a group of metatheoretical results, the theorems of formal limitation may be described in recursive terms: The metatheoretical reasoning they rely upon involves a dependency upon the possibility of perpetually extending a line of reasoning.

What characterizes the constructive [i.e., constructive mathematics] is the possibility of going always further, which is therefore the notion of “*after*.”

The original field of application of this notion is that of the integers. The series of integers can indeed be characterized by means of the notion of *successor*, and we know that solely by means of the successor operator we can constitute the whole series of integers beginning with 0.

But the series of integers does not exhaust the field of application of this notion.... (p. 434)

Let us return to the conceptual context of this study: In connection with the in-principle extendability of the contents of many contexts of reference, purely Leibniz boundaries (see {14}) may be involved, which means that such extendability is an expression simply of an expectation, based on a rule or a habit that has been formed in similar instances. Alternatively, such extendability may involve an attempt projectively to claim that such extendability has the putative meaning of a series “already completed,” of a series whose members are projectively considered to possess an existence autonomous of the referential framework which must be presupposed for reference to that series to be possible. We encountered such extendability in a variety of referential contexts, one of which was in the discussion of time-order ({22.1}).

We have observed already, albeit here in greatly abbreviated form, that the theorems of formal limitation rely upon metamathematical reasoning combined with the application of recursive procedures which, on a theoretically fundamental level, involve the recurrent application of the notion of “the next” in a potentially extendable series of elements. The formalizability of a system rests upon how such recursively applied reasoning in the theory of formal systems is to be understood. Such reasoning has often been understood by mathematicians, in particular intuitionists such as L. E. J. Brouwer and Arend Heyting, to have a temporal structure which we find in an in-principle continuable series of successive elements.<sup>365</sup>

We have noted that the temporal nature of identifying reference, the basis for the successive reiterability of reference, has a fundamentally recursive character, and yet the successive reiterability of reference—the in-principle extendability of reference with respect to a given referential field—is subject to the metalogical constraints recognized by the metalogic of reference. Those constraints, as we’ve seen, in particular take the form of reactive horizons which, when claims are made that seek putatively to transgress those horizons, short-circuit the possibility of reference and meaning.

At this point, we need to note that the understanding of the metamathematical reasoning entailed by the theorems of formal limitation diverges from the reasoning employed by the metalogic of reference. To make this distinction clear, I again quote Ladrière, whose philosophical background in phenomenology supported his recognition of the temporal structure of the mathematical constructive reasoning that is relied upon by the theorems of formal limitation:

---

<sup>365</sup> In this connection, we recall (see {19.2}) Brouwer’s “intuition of two-oneness” which refers to the temporally successive nature of human consciousness that underlies the experience of counting and the concept of number.

[T]he very structure of temporality excludes the possibility of a total reflection; there is no moment that would itself be a recapitulation of all the others, a present that would absorb within itself the past as it would the future and which would expand to the dimensions of the whole experience. The only present that is accessible to us is a present marked by precariousness; rocking without cease in the non-actuality of that which is no longer, it is at the same time always open to the non-actuality of that which is not yet.... (p. 438)

The meaning of this informal, poetically expressed passage, unusual in Ladrière's highly technical work, requires some brief interpretation and commentary. The latent argument seems to me to be that the temporal ordering of mathematical experience, of the constructive mathematician's consciousness, makes "total reflection" impossible because, in Ladrière's view, there putatively exist autonomous moments in time that cannot be included in the experience of the present; some of these are past moments, some relate to times in the future. Yet, this interpretation is not consistent with Ladrière's conjoined claim when he says that these moments, which cannot be included in present experience, are "non-actual." Let us suppose that in a contradiction-tolerant logic both claims are made: both the claim of independent existence (of "all the other" moments that cannot be included in present experience), alongside the claim of their non-existence (of what "is no longer" and "is not yet").

Both of these claims, I think the assiduous reader of the present study will immediately see, are projective: Both attempt to refer beyond the metalogical horizon of the constructive mathematician's consciousness, the context Ladrière has in view. Neither claim can, in principle, possess possible meaning.

From a phenomenological standpoint, Ladrière recognized that the constructive nature of mathematical reasoning is fundamentally temporal in nature. Because he understood mathematical reasoning in this light, one of the principal philosophical conclusions which he drew from a study of the theorems of formal limitation he chose to express in hypothetical form: "If it is true that the structure of the constructive symbolizes that of temporality, the impossibility of a total system combines the impossibility of total reflection" (p. 443).

In contrast, the realization to which we are brought by the limitative results of mathematical formalization—insofar as they apply to the formal structure of the metalogic of reference—requires de-projective analysis. As I

understand the philosophical ramifications of the internal limitations of formalization, that realization should, like Ladrière's conclusion, be expressed in hypothetical form:

If (i) the metatheoretical framework of the metalogic of reference is, as this study has claimed, capable of complete reflexivity; if (ii) the metalogic of reference, as a theory developed on the level of maximum theoretical generality, is its own metatheory; and if (iii) the recursive, temporal structure of the metalogic of reference comprises a self-enclosed theory, as developed in {4.10}, then its logical structure may not be deductively formalizable.

I have deliberately chosen to express this realization hypothetically, leaving its conclusion open to revision. One reason for this is developed in the next section; a second reason is that it may prove to be the case that our understanding and our means of mathematical formalization may eventually evolve in ways that cannot today be anticipated from the standpoint which the theorems of formal limitation together define.

### **§3. The formalization of complex or even ordinary reasoning**

Earlier in §1.2 of this Supplement, I raised the question whether ordinary, everyday reasoning, as well as the more complex, multi-leveled informal reasoning that is indispensable to much philosophical thought, may—perhaps in principle—be unformalizable. This question clearly poses the question of formalization from a different perspective than do the theorems that demonstrate the *internal* limitations of formalization. To raise the question of the formalizability of everyday and more sophisticated reasoning is to ask a question that approaches the limits of formalization *externally*, from outside of the structure of formalized systems, beginning from the standpoint of human reasoning itself. Since the objective of this Supplement is to describe, in a manner yet to be made clear, the formal structure of the metalogic of reference, both the internal and the external questions regarding its capacity to be formalized need to be considered. The preceding sections have pointed strongly in the direction of the conclusion that the logical structure of the metalogic of reference may not be formalizable given the internal limitations of deductive formalization itself. Although the case against the deductive formalizability of the metalogic of reference is strong, the case would be stronger still should we be forced to give a similar answer based on an appraisal of the nature of human reasoning.

It is important when speaking of the “formalization” of human reasoning to make several distinctions: The variety of formalization that has been at issue in connection with the theorems of formal limitation is of course



deductive formalization, the variety familiar to mathematicians and logicians that results when an artificial symbolic language is used to define a system of deduction, commonly axiomatic or a system of natural deduction. However, in addition to deductive formalization, it is also sometimes possible to represent reasoning procedures *algorithmically*, often in a manner that lends itself to expression in computer code. There is also the less strictly regimented approach to the representation of reasoning procedures as found in *expert systems research* which commonly emphasizes the formulation of a network of if-then production rules rather than necessarily programmable procedural code. Related to the approach of expert systems, there is *symbolic artificial intelligence* whose objective is symbolic language-based formulations of problems, search and solution routines, etc. These as well as other related approaches may be grouped together and for the purposes of simplicity may loosely be called ‘*non-deductive formalizations*’: By the use of this phrase we shall have in view all such approaches that seek to characterize and reproduce human reasoning processes by means of *representations of the formal structure* of those processes *in language*, usually employing specialized artificial languages, including computer code.

In addition to such non-deductive means of representing human reasoning processes, we need also to recognize attempts, not to describe and reproduce those processes in a language, but rather to *simulate* or *replicate* human reasoning processes *in physical or in electronic form*: The evolution of artificial and bioengineered neural networks is an example. Whether such simulation and replication research is significantly successful as it might relate to the reflective reasoning processes required by the metalogic of reference, remains to be seen. A discussion of this potential must fall outside of the scope of this book.

Whether human reasoning, of the everyday or the conceptually complex variety, can, in principle, be deductively formalized—or non-deductively formalized in any of the above senses—is as yet an unanswered question. Although the question is clearly of philosophical interest, answering that question has, by consensus over a period of many decades, in the main been delegated to researchers in computer science and in particular to researchers in artificial intelligence.

One of the earliest cognitive scientists and AI researcher who was motivated to give free expression to his *beliefs* about a likely answer to this question was Marvin Lee Minsky (1927–2016). In a short and now well-known monograph, Minsky (1974) presented an early set of suggestions which he argued could be used as a paradigm for the machine simulation of human thought processes. In an appendix, “Criticism of the Logistic Approach,” he

stated, but did not prove, his answer to the more limited question whether human reasoning can be deductively formalized. If his observations are extended, as I wish to do, in an answer to the wider question of non-deductive formalization, his remarks pose an informal but articulate challenge to the successful representation of human reasoning by a formalized system, or, by extension, by means of natural or artificial languages. Let us see why this is so.

During the half-century that has elapsed since Minsky expressed his remarks concerning the shortcomings of a logistic approach, they have been much discussed, but no definitive response to the concerns he expressed has yet been found. Here is a condensed account of his position, in his own words:

There have been serious attempts, from as far back as Aristotle, to represent commonsense reasoning by a “logistic” system.... I think such attempts will continue to fail, because of the character of logistic in general rather than from defects of particular formalisms....

Even if we formulate relevancy restrictions, logistic systems have a problem in using them. In any logistic system, all the axioms are necessarily “permissive”—they all help to permit new inferences to be drawn. Each added axiom means more theorems, none can disappear. *There simply is no direct way to add information to tell such [a] system about kinds of conclusions that should not be drawn!...*

Why then do workers try to make logistic systems do the job? A valid reason is that the systems have an attractive simple elegance; if they worked this would be fine. An invalid reason is more often offered: that such systems have a mathematical virtue because they are

- (1) Complete—“All true statements can be proven”; and
- (2) Consistent—“No false statements can be proven.”

It seems not often realized that Completeness is no rare prize. It is a trivial consequence of any exhaustive search procedure, and any system can be “completed” by adjoining to it any other complete system and interlacing the computational steps. Consistency is more refined; it requires one’s axioms to imply no contradictions. But *I do not believe that consistency is necessary or even desirable in a developing intelligent system.* No one is ever completely consistent. *What is*

*important is how one handles paradox or conflict, how one learns from mistakes, how one turns aside from suspected inconsistencies....*

“Logical” reasoning is not flexible enough to serve as a basis for thinking; I prefer to think of it as *a collection of heuristic methods*, effective only when applied to starkly simplified schematic plans. The Consistency that Logic demands is not otherwise usually available—and *probably not even desirable*—because consistent systems are likely to be too “weak”....

I cannot state strongly enough my conviction that the preoccupation with Consistency, so valuable for Mathematical Logic, has been incredibly destructive to those working on models of mind. At the popular level it has produced a weird conception of the potential capabilities of machines in general. At the “logical” level it has blocked efforts to represent ordinary knowledge, by presenting an unreachable image of a corpus of context-free “truths” that can stand separately by themselves. (pp. 74-78, italics added)

Let us detach and re-phrase from these passages four bold claims that Minsky has made unflinchingly, and then translate them into a broader context with the formal structure of the metalogic of reference in view:

- (a) There is “no direct way” to inform logical systems about “kinds of conclusions that should *not* be drawn.”
- (b) Undue emphasis should not be placed on logical completeness, since there may be other ways in which true statements can be recognized.
- (c) Logical consistency is not necessary or at times even desirable.
- (d) The effective use of logical reasoning is essentially heuristic.

If (a) were to be the case as applied to the metalogic of reference, then we may conjecture that it may be difficult, and, by extension, perhaps impossible, to include in a rigorously expressed deductively formalized system *or* in a non-deductive system statable in language any set of rules that will effectively *exclude* metalogical projections. This limitation is clearly conjectural, but at this time, it cannot categorically be ruled out.

In connection with the theorems of formal limitations, we have noted that

they share in common a reliance upon metamathematical reasoning. This reliance itself is confirmation of (b): Metatheoretical reasoning exemplifies one of the “other ways” in which true statements, undecidable in a given formalized system, may nonetheless be recognized as true.

In connection with (c), a reflective analysis of metalogical projections has brought us face-to-face with the widespread human predisposition or compulsion to transgress the horizons of referential frameworks. The problem of putative meaning ( $\{12\}$ ), of recognizing the delusional nature of purportedly meaningful projections, can, as we’ve seen, be resolved in a clear way when one takes the reflective, analytical trouble to do this. And yet it would be unrealistic to overlook the fact that many people, including many philosophers, do indeed fall victims to projective delusion. The logical structure of the metalogic of reference seeks to bring about the consistency of concepts, claims, positions, and theories with the referential preconditions necessary for their possibility and meaning. This, as I have emphasized, is a purely conceptual undertaking. However, the pragmatical application of the metalogic of reference to human reasoners can, as previously observed, be a logically messy affair.<sup>366</sup> Whether the conceptual or the pragmatical structure of the metalogic of reference can be deductively or non-deductively formalized in a manner provably consistent is not a question I will be especially concerned with in subsequent sections. As Minsky suggested, excessive concern over the

---

<sup>366</sup> It is perhaps a bit messier than most philosophers prefer to think: Human experience is surely varied and complex in structure, so much so that one may conjecture that any attempted formalization of its organizing principles may lead us to suspect that such a formalization may not be altogether coherent, and perhaps cannot, in principle, be made to be coherent. We have some indirect evidence for the fact that the organizing structure of human experience does and will at times and in certain individuals malfunction (as in neurosis), break down (as in psychosis), or become non-functional (as in catalepsy and brain death).

Though this is the case, theorists of reference tend to ignore such dysfunctions, and they direct attention instead to the relatively stable portions of human experience that are apparently, at least to the more casual analyst’s eye, free from the potential of structural short-circuiting. We seldom make explicit that this restriction of interest—to what is more readily analyzed in coherent terms—expresses a deliberate, though usually unexamined, preference. Yet preference it is. We prefer to study what is most amenable to study, given the conceptual tools at our disposal. And this is surely sensible and pragmatic.

But we should not, in the process, lose sight of the fundamental restriction in the scope of what we wish to study, brought about by the limitations of our preferences. When we are reminded of this, we may be persuaded to accept the doubt that stems from the *perhaps inescapable* limitation that a genuinely general metalogic of reference, if it could be expressed as a deductively formalized system, may not be coherent in the sense of always avoiding contradiction—the logical equivalent of an electrical/neural short-circuit. We might, furthermore, be inclined to ask whether the very structure that permits such short-circuiting is a structure that is necessary for total reflexivity.

unavailability of consistency proofs can in some contexts hamper constructive efforts.

...

We have so far discussed some of the chief advantages as well as shortcomings of deductive formalization. This was followed by a description of internal limitations of formalization that may plausibly be applicable to the formal structure of the metalogic of reference. Thus far, we have considered the following limitations:

- (i) the impossibility of comprehensive or total formalization
- (ii) the impossibility of the total reflexivity of formalized systems
- (iii) the possibility that the reflexive nature of the metalogic of reference, its development on the level of maximum theoretical generality, and its recursive, temporal structure as a self-enclosed theory, may preclude, in principle, its deductive formalizability
- (iv) the possibility that complex or even ordinary human reasoning may not be represented and reproduced successfully in non-deductive artificial or natural language

As we have seen, there are theoretically compelling reasons to accept the first two of these. I have expressed (iii) and (iv) hypothetically, leaving open whether the logical structure of the metalogic of reference may find an adequate formalized expression, either through unanticipated extensions of what we now accept as the scope of deductive formalizability, or through non-deductive means of characterizing and reproducing its structure algorithmically, employing an expert systems approach, in computer code, or other related methods. It would require a very considerable extension of this study to attempt to provide more complete answers to the questions these limitations together pose. In this study, I have chosen to leave these questions unanswered, primarily because not possessing answers to them does not significantly affect this study's objectives, which are, I think, more suitably understood in terms of the heuristic nature of the metalogic of reference.

#### **§4. The formal structure of the heuristic method of the metalogic of reference**

Previous chapters in this study have stressed the heuristic nature of the metalogic of reference. Readers will recall that the principles that were formalized

in {11} were characterized there as heuristic tools of analysis. In the metalogic of reference, those principles serve as metatheoretical rules. As noted in {11.4}, we have not been concerned with whether these heuristic principles are logically independent of one another, or in developing a completely formalized deductive system. In {14.4}, the rationale was given for using such a heuristic approach: First, as a set of heuristic tools, they supply guidelines for the solution of a problem—for the metalogic of reference, the central problem is the identification, correction, and elimination of projective concepts, claims, positions, and theories. Second, a flexible heuristic approach is best-suited to a discipline like philosophy in which contention and the taking of positions often dominate: A heuristic approach can more easily accommodate and be applied to the diversity of philosophical frames of reference, something which a rigid, strict algorithmic formulation, for example, is less able successfully to accomplish because of the need for its shared acceptance among philosophers. Third, as we have seen in this Supplement, there are good—some would even say compelling—reasons to think that the structure of the metalogic of reference, given its defining properties, does not, perhaps in principle, lend itself to deductive or even to what I have called ‘non-deductive formalization’. If this is the case, then it makes sense that its approach should not be claimed to be more than a set of heuristic tools that can be of assistance in detecting metalogical horizons and in constraining delusion-inducing attempts to transgress those horizons. As a heuristic method, one which at the present time is not known to be either deductively or non-deductively formalizable, de-projection must rely upon a human metalogical analyst’s reflective skills and his or her ability to render explicit a concept’s, a claim’s, a position’s, or a theory’s fundamental referential structure—the referential preconditions that must be satisfied in order, in principle, for that concept, claim, position, or theory to refer and to have meaning. These at present exclusively human conceptual skills and ability are described in detail in Appendix II. In this sense, because of the essential reliance upon the reflective analytical skills of the analyst, the formalized expression of the metalogic of reference does not aim, as do many formalized systems, to be self-contained.

In the sections of this Supplement that follow, to make the formal structure of the heuristic method of the metalogic of reference as clear and explicit as possible, I have chosen a format that might be called an ‘*axiomatized heuristic*’. The format is axiomatic, but axiomatic on the surface only since there is no special concern, as mentioned previously, to show that each of the axioms is logically independent and cannot be dispensed with, or to show that the set of axioms is minimal in number, or to deliberate over the issues of consistency or completeness that pertain to a formalized system. Such features

of many axiomatic formalizations of course provide elegance, but they can also sometimes interfere with ease of understanding, and ease of understanding is my principal purpose in what follows.

The axiomatized formulation is, with a few exceptions which are noted, intended to be read separately from the somewhat more complex and intricate formalization in {11}.<sup>367</sup> In this Supplement, the deductively structured representation of certain of the main principles that define the logical structure of the metalogic of reference is intended as a potentially informative alternative to that found in {11}. The axiomatic formulation should not be regarded as providing more than a simplified *formal sketch* of the heuristic method. Like any sketch, it is at times recognizably and unavoidably over-simplified and imperfect. Nevertheless, a sketch of this kind may better serve the needs and interests of some readers.

### §5. Informal preliminary

As a preface to their formalized expression, this section briefly recapitulates in a simplified and informal way a group of concepts and principles fundamental to the metalogic of reference, supplying occasional examples.

As we have seen, a chief purpose of the metalogic of reference (called ‘MoR’ in its representation as an axiomatized heuristic) is to provide a negative definition of the domain of possible sense through the detection of metalogical horizons ({14}). As the formal expression of a metalogical theory, MoR represents a methodological meta-framework for the analysis of the referential preconditions of individual theories (or of the referential preconditions of individual referring concepts, claims, or positions). The now-familiar, central objective of that meta-framework is to study the general preconditions of reference, possibility, and meaning. In this study, we have come to understand preconditions of possibility and meaning in terms of parameters of referential constraint.

The defining interest of MoR is exclusively in these general constraints on reference, and not in any individual theoretical identification framework; as a result, rules or conventions that are framework-specific have been left open. MoR seeks to give a formalized and schematic expression of this book’s method of studying those preconditions which function as referential constraints for any particular identification framework, constraints that must be observed in order to avoid metalogical projection ({13, 14}).

---

<sup>367</sup> Portions of subsequent sections of this Supplement are based on, albeit with various changes, Bartlett (1975). The set of axioms is generally the set on which von Weizsäcker commented in the Foreword to this book.

The following are basic concepts of MoR:

*Identification framework*

A theory (or concept, claim, or position) may be analyzed in terms of the “identification framework” it presupposes. An identification framework is a system or structure  $S$  that satisfies the following conditions:

- (a) there is a finite, indeterminate, or potentially infinite number of elements which may themselves be sets of elements, sets of sets, etc., such that  $S$  provides a basis for identifying reference to these elements;
- (b) identifying reference to an element in  $S$  identifies that element so that it can be re-identified; thus
- (c)  $S$  is essentially a temporal structure in that it provides a basis for identifying reference to an element subsequently re-identified as “the same element.”

The structure of a theory is further determined when to (a) – (c) are added particular constraints limiting the range of a given theory’s possible reference. These parameters prescribing a theory’s range of possible reference determine the essential referential structure of the theory. Another way of expressing this is that the referential structure of a theory is a function of the ensemble of objects to which it permits reference. For MoR, such objects will normally remain completely undetermined “*particulars*.”

*Identifying reference*

An identifying reference is such that an ascription to that which can be the object of an ascription establishes that what is ascribed and that to which ascription is made are one and the same.

*Examples.* Identifying reference to a point in 3-space is made by  $(0,1,\sqrt{2})$  where the three coordinates together assign or ascribe a location to a point within an appropriate coordinate frame. The concept of a coordinate system illustrates identifying reference in an ideally simple sense since the position determined by a set of coordinates will automatically be satisfied by the point they define.



In the case of more complex and less easily formalized contexts—for example that of persons in a certain situation which has, among others, cultural, spatial, and temporal parameters—identifying reference to an object of an ascription is often much less precisely determined than in the above example and is likely to be a function of the shared “understanding of the situation” of the persons involved.

Identifying reference is universally relied upon, whether in the sense of uniquely identifying an object, or it may be in a rule-determined or vague manner. The concept of identifying reference is of fundamental importance in any theoretical investigation. From the standpoint of MoR, identifying reference is a presupposed, primitive concept.<sup>368</sup>

### *Particulars*

A particular is defined as the object of an identifying reference. This restricts the use of the concept of particular to objects with an identity which in principle it is possible in some manner to characterize. This definitional restriction is later transformed into a fundamental premise of MoR.<sup>369</sup>

It follows that the “preconditions of possibility” of a particular are co-extensive with the referential preconditions of the identification framework permitting reference to that particular. As distinct referential preconditions are involved, distinct ensembles of particulars are defined.

In general, a particular comprises some structurally definite organization, a context, or a situation which can be identifyingly referred to. A matrix, for example, consisting of one or more possible subjects of identifying reference unified according to rule in terms of order relations itself comprises a possible object of identifying reference and may be understood in terms of the mathematical concept of function. In short, we recognize a complex variety of possible objects of identifying reference which are grouped together under the

---

<sup>368</sup> In a somewhat parallel way, Paul Lorenzen calls unambiguous descriptive identifications ‘*elementary*’: The “decision to accept elementary ways of speaking is not a matter of argument. It does not make sense to ask for an ‘explanation’, or to ask for a ‘reason’. For to ‘ask’ for such things demands a much more complicated use of language than the use of elementary sentences itself. If you ask such questions, in other words, you have already accepted at least the use of elementary sentences” (Lorenzen, 1969, p. 14).

<sup>369</sup> We recall Quine’s related dictum “no entity without identity” (Linsky, 1967, p. 27) as well as Wittgenstein’s remark that “It is impossible to represent in language anything that ‘contradicts logic’ as it is in geometry ... to give coordinates of a point that does not exist” (Wittgenstein, 1961, 3.035). Cf. also Husserl (1929, §65).

general concept of particular. The term ‘particular’ is therefore not limited to “ultimate simples.”

It is important to note, however, that the concept of particular does not presuppose structures of the subject-predicate type insofar as these involve the predication of properties to a grammatical subject. The language employed in defining ‘identifying reference’ should not be misunderstood as involving “ascription of properties to a grammatical subject.” (For extended analyses, see the main text, e.g., {24.2} and *passim*.) A particular is the referential foundation for the possibility of such predication; the reverse will not always be the case since not all contexts of reference permit predication to a grammatical subject.<sup>370</sup>

Thus, a particular is not only an individual object of reference, but may be thought of as a general form of *functional marker*.

*Example.* In the expression ‘ $f(x)$ ’, ‘ $f$ ’ may be called such a functional marker. The role which it plays, in relation to the  $x$ -variable, is that of a constant which fixes conditions according to which  $x$  may vary. If  $f(x)$  and  $g(x)$  are both specific functions, then two (perhaps related) functions, or particulars, are considered:  $x$  is a functional variable which may take certain values; similarly for  $y$ . The two functional markers,  $f$  and  $g$ , are respectively self-identical and constant, “the kernels...which...remain within an undefined generality as a something...identical over identifications” (Husserl, 1929, §43).

The elements which may satisfy the functional variable clearly depend upon the particular in question. The individualizations permitted establish the specific properties of what otherwise is often a highly general referential structure. Since the “objects” for the purposes of MoR remain completely undetermined, the concept of particular is a concept of a purely abstract referential structure.

A more concrete, subordinate illustration follows.

*Example.* Consider a number of images perceived sequentially on a screen. In the case that general relations between

---

<sup>370</sup> Compare, for example, “feature-placing languages” in (Strawson, 1959, Part II), some Wittgensteinian language-games, or process-oriented natural languages in which the grammatical function of predication plays a greatly weakened role.

individual images or groups of images are perceptible to a viewer, these relations may enable him or her to comprehend the sequence, or parts of it, as unitary, perhaps as referring to a “particular” which the group of images functions to identify. Roughly speaking, such a particular object of reference becomes intelligible to a viewer literally as a function of the individual images seen successively.<sup>371</sup> A particular is thus equivalent neither to the succession of viewed images nor to the perceived relations (e.g., of similarity) between individual images, but rather to the referential structure as a whole which relates the sequence of images—or parts of that sequence, *via* relations between individual images—to the identity of an object which in retrospect is recognized as their unitary theme.

To summarize, the concept of particular is fundamentally parametric, involving as it does the specification of one or more identity conditions and a field of variability within which possible values are related according to one or more regulative principles.<sup>372</sup> An identity condition is, generally speaking, a relation constitutive of a variation pattern or rule of order, and may be thought of as a differentiation within the range of variation of a parameter of higher order. The identity condition restricts and regulates this range of variability.

*Example.* In the previous example, a succession of images provided a context permitting identifying reference to a particular which in retrospect is perceived to provide a principle of unification (a “variation pattern” or “rule of order”) of the sequence, or parts of it. However, *the previous example itself*—as an object of identifying reference—is situated within a larger context which permits referential discrimination and eventual identification of various particulars, among them the example in question. This larger context—for instance, that provided by the experiences we remember having had in the hours immediately preceding our viewing of the sequence of images—makes it possible to refer to the succession of

---

<sup>371</sup> We note the connection between the identification of particulars and pattern recognition.

<sup>372</sup> As Scott Buchanan expressed this: “The central principle of the notion of parameter is variability limited and controlled by identity conditions.... Such identity conditions have been called conditions of possibility..., principles of synthesis for a manifold or multiplicity of particulars” (Buchanan, 1927, p. 48).

images and the relations these exhibit in that succession. And this larger context is in turn situated within yet a larger one, perhaps the context presupposed when, in reading, we reconsider the previous example.

### *Descriptions*

The content of what is ascribed in an identifying reference comprises a “description” of the object to which there is reference. Depending on the context of reference, what is thereby described permits a relatively unambiguous differentiation within the range of variability of the context.

*Example.* A set of instructions, a map, a compass, and a person competent to interrelate these appropriately, together establish a context of reference in terms of which the set of instructions constitutes an explicit description in the sense the instructions intend, e.g., to specify a geographical location.

As the concept of description is understood, then, a description presupposes an appropriate identification framework since a description itself comprises an object of reference and is thus a particular.

As noted in {10}, identification may be conceived as involving a descriptive component of specification and a component of satisfaction when what is descriptively specified coincides with that to which there is reference. Together these two components provide the basis for the identification of, or identifying reference to, a particular. The two components make possible identifying reference to an object of reference such that what is ascribed and what is specified are identical, constituting the identity of the object in the given context of reference.<sup>373</sup>

## **§6. Deductive representation**

Before proceeding, it will be helpful to describe the general character of the formalized representation of MoR.

In what follows, a number of concepts basic to the metalogic of reference, discussed informally in the preceding section, are expressed in symbolic form.

---

<sup>373</sup> A familiar example was mentioned in {3.3}, that of a single expression that performs both identifying and descriptive functions together: the coding system employed by the U.S. Social Security Administration, according to which a given Social Security number serves both to identify an individual and to provide a description of certain of that individual’s properties.

To an extent the resulting formalization expresses more perspicuously and compactly what in English can be cumbersome and lengthy. In this, it is important to re-emphasize that recourse here to the elements of deductive theory is not to be taken to imply that either a complete or an adequate translation of MoR into deductive form is possible. The question whether this can in principle be accomplished comprised the subject-matter of the preceding §§2–3. MoR is presented in deductive form but without the normally implicit claim that it constitutes a deductive formalized system. It does not. It rather comprises an “axiomatized heuristic.” Use of a deductive representation is not therefore intended to be suggestive of the adequate or the complete formalizability of MoR, but rather to express, perhaps more clearly than can be done in natural language, certain of the fundamental and elementary principles and relations between basic concepts which are important in analyses of the metalogical preconditions of reference.

*Elements of Deductive Theory*

The main elements of deductive theory are generally taken to include the following:

	<u>Symbolized by</u>
Primitive concepts	C
Rules of derivation (postulates of inference)	P
Definitions	D
Axioms	A
Theorems	T
Interpretations	I

A deductive theory may be divided into

- 1 a set of elements, sets of sets of elements, sets of sets of sets of elements, etc. (e.g., in a deductively formulated physical theory, the set of events),
- 2 a set of relations which hold between elements, sets of elements, etc. (e.g., a time-relation between physical events), and
- 3 a set of propositions that specify the properties of these relations (e.g., physical laws).

The sets 1 and 2 are used to formulate the set of *concepts* basic to the deductive theory, while the set 3 is usually thought of as formulating *laws* to which the concepts conform.

To develop a deductive theory, one requires: a set of *symbols* which make up the available symbolic vocabulary and which can be used to express the *primitive concepts* of the theory; a set of *rules of formation* which indicate how permissible, well-formed strings of symbols are to be constructed; a set of propositions about these symbols which formulate *primitive rules of derivation* in a manner so that the meaning of a primitive concept is contextually expressed by the rules of derivation; an initial group of well-formed strings called '*axioms*'; frequently, because this is convenient, *definitions* which provide a means for replacing longer strings by shorter ones; *theorems*, which follow from the application of the rules of derivation to the axioms; and finally and where this is useful, a set of statements making up an *interpretation*, establishing a correspondence between the deductive system and a system or framework of another sort which serves to exemplify the structure of the former.

Since the metalogic of reference is a metatheory concerning the nature of theory, I mentioned earlier that we suspend any interest in special classes of *objects*, and instead work with a neutral system of concepts in which only undetermined *particulars* are studied. (While we do this, we can of course bear in mind the ease with which one can shift out of this way of representing things and return to particular sorts of "objects" in which our interest may lie, which are no more than specific kinds of objects of reference, sets of these, etc.) This suspension of interest in particular varieties of objects responds to a philosophical need for an approach that aims first for simplicity and generality, and then for systematic applicability to more specific problems.

On behalf of simplification, I have sacrificed elegance and technical sophistication in the following formalized representation, and will not burden the text with a complete commentary on the notation and rules of derivation that are used, as would be the case were MoR to be advanced as a completely formalizable system. The formalized representation generally employs rules of derivation that are familiar to any reader of texts of mathematical logic.<sup>374</sup> Some basic set theory is also assumed. Rules of derivation specific to MoR that may be unfamiliar to many readers are those designated by 'P' ("postulates of inference"). Postulates are selected for their heuristic usefulness in referential analyses and not for their independence from one another. Ele-

---

<sup>374</sup> For the format used in proofs and for many of the rules of derivation, cf. Lemmon (1971/1965) and Fitch (1952). For a general description of the axiomatic method in mathematics, see, e.g., Kleene (1952).

ments of the deductive representation of MoR are briefly discussed when they do not derive from the general literature.

MoR departs from familiar approaches to formalization in a number of important ways which deserve comment: The formalized representation of MoR is unconventional and indeed unusual in a number of respects, some of which we have already noted:

(1) I stated in the previous section that the formalization given in this Supplement is a “sketch”; by that I mean not only that it is a simplified formal representation, but that the formalization that is given in this Supplement cannot, due to constraints of space, be provided in a form that will supply an exacting reader with all of the deductive means that would normally be identified explicitly. To supply a more adequate formalization would require a separate book in itself. Without a doubt, the demands of the subject-matter exceed what can be accomplished in the space available.

(2) The formal representation of MoR is a *hybrid* formalization: On the one hand, its logical structure is presented axiomatically and a basic set of rules of derivation is given in the form of postulates as is common in axiomatic presentations; on the other hand, readers will find that a set of natural deduction rules of derivation is also incorporated. By means of this hybrid combination, the objective has been to exhibit the reasoning in as close to a natural manner as possible.

(3) In light of the preceding discussion of the internal limitations of deductive formalization, and given the reflexive nature of the metalogic of reference as a theory on the level of maximum theoretical generality, an attempt to provide MoR with customary formal means to express what is ordinarily regarded as a hierarchy of levels of language is not considered inappropriate. As an axiomatized heuristic, the formalism that is employed continuously presupposes a reflective, reflexive meta-framework on the part of the reader, who is called upon to provide the multi-leveled referential context necessary to understand the dynamics of MoR’s formal representation. In this connection, the symbol for metatheoretical derivation or metalogical entailment, ‘ $\Vdash$ ’, defined in {11.4}, will be found in a number of theorems and one postulate. Use of ‘ $\Vdash$ ’ should be thought of not only in terms of its defined meaning, but as an explicit reminder that an entailment relation is metalogical and requires reliance upon metatheoretical reasoning.

(4) Most individual proofs in MoR use a format that consists of an abstract of a series of inferences. The format is generally condensed, for two reasons: for simplicity—to place in relief some of the important deductions and inference patterns of MoR in a manner so that these do not get lost in an over-emphasis of proof details—and for economy of space. Given the reliance

upon the reflective, analytical resources of readers presupposed by (3), the use of proof abstracts is appropriate and should make good sense; most of the inferences require only elementary reasoning.

(5) A fifth and relatively uncommon feature of MoR is the use of arbitrary names. For readers unfamiliar with the approach advocated by logician E. J. Lemmon (1971/1965), the role of arbitrary names requires explanation. That role in many works in mathematical logic is played by free variables; instead, MoR often employs arbitrary names ('*a*', '*b*', '*c*', ...). If it is the case for an *arbitrarily selected* object that  $Fa$ , then it is—subject to a strict restriction—valid to conclude  $(x)Fx$ . That is, if a *truly arbitrarily selected* object has a certain property, then everything does. The restriction is that the arbitrarily selected object be truly *arbitrary* and *typical* in nature: the inference from  $Fa$  to  $(x)Fx$  can rest on no special assumptions about  $a$ . For a full account of this restriction along with examples of its use in proofs, see Lemmon (1971/1965, pp. 107-109, 116, and *passim*). In addition, if it is the case that  $(x)Fx$ , then, for an arbitrarily chosen object  $a$ , it of course follows that  $Fa$ , and it also follows that  $Fm$  for an object with the proper name '*m*'.

A word should be said about a technique of proof that is employed in connection with certain of the rules of derivation and axioms. Many people tend—often with good reason—to regard an axiomatic formulation as fundamentally unconvincing since the acceptance of specific rules of derivation and axioms can sometimes seem unnatural and arbitrarily stipulative. To respond to this potential concern and to demonstrate the strength of the basis of MoR, a proof technique is used which in {16} was named 'self-validation'. In this Supplement, a rule of derivation or axiom *self-validates* if its denial results in projection, i.e., in metalogically self-undermining inconsistency ({13}); for the formal definition of self-validation, see {11.4}. The technique of self-validation makes it possible to utilize the metalogical presuppositions entailed by the denial of an assertion to justify the assertion. The technique is somewhat similar to the dialogue-operative logical proof technique advanced by Paul Lorenzen, in which proof strategies reflect the *de facto* shared rhetorical basis of disputants relative to a universal audience.<sup>375</sup> The important difference is that Lorenzen's proof strategies relate to what is *in fact* the case, whereas the technique of self-validation relates to what *must, in principle*, be the case.<sup>376</sup>

<sup>375</sup> See Lorenzen (1969) and (1969a), as well as Perelman & Tychka (1971).

<sup>376</sup> The principal difference between "self-validation" and an "operative logic" lies in the following: For Lorenzen, appeal to a common basis of argumentation is justified because *there is, in fact*, for a group of disputants, no appeal to an alternative basis. In the metalogic of reference, the appeal in self-validation is justified on a different level—namely, in terms of pre-



In light of the preceding informal discussion, we reiterate the definitions of particulars and identifying reference given in {11.4}:

D 1            A *particular* is a possible object of identifying reference

which presupposes the concept of identifying reference, and the stipulative definition of identifying reference:

D 2            An *identifying reference* is such that an ascription to that which can be the object of an ascription (namely, a particular) establishes that what is ascribed (one or more properties, relations, a description, etc.) and that that to which the ascription is made are one and the same (identification)

which presupposes the concepts of particular, description, and identification, and leaves the concept of “ascription” as conspicuously retaining, again in an unavoidably circular way, the concept of identifying reference which the formulation of (D 2) must presuppose.

For economy, MoR retains D 1, permitting the concept of identifying reference to an object to be logically primitive, and transforms D 2 into a later interpretation (I 18). In D 1, when we grant the concept of identifying reference the status of a logically primitive concept—that is, one that remains formally undefined by means of other logical operations in MoR—D 1 evidently brings with it the concept of possibility (“a particular is a *possible* object of identifying reference”). We are of course free also to include possibility under the veil of logical primitiveness; nonetheless, for the heuristic purposes of MoR, readers will want to keep in mind the understanding we have reached of the general concept of possibility in terms of a reference system’s parameters of referential constraint. As will be recalled from {7.5}, the concept of possibility is understood in terms of what is permitted by a system of referential constraints, or equivalently, what is not prohibited by those constraints. That is,  $\diamond p \equiv \sim \square \sim p$ ,  $p$  is possible (i.e.,  $p$  is referentially permitted) if and only if it is not necessarily the case that not- $p$  (i.e.,  $p$  is not referentially prohibited). Additionally, what is necessary is conceptually equivalent to its being impossible for it not to be the case, that is,  $\square p \equiv \sim \diamond \sim p$ ,  $p$  is necessary if and only if

---

conditions of *possible* reference insofar as there *can be* no appeal to an alternative basis.

it is impossible that not- $p$ . In this sense,  $\square$  is not logically primitive, but is defined in terms of  $\sim$  and  $\diamond$ .

Accepting the logical primitiveness of certain concepts in a given formalization should present no difficulty since it is well-known that, as concepts are used to explicate concepts, an indefinitely continuable process results if one tries to explicate *every* concept. The moral is not to try and to accept as primitive certain concepts, with identifying reference and possibility as both central and unavoidable, and thus logically primitive in MoR. As is tautologically the case, one must have a starting point in order to begin.

We may now begin the formal representation of MoR:

C 1                     $\mathcal{R}$

I 1             $\mathcal{R}$  is the class of identifying references. A member of  $\mathcal{R}$  is called ‘R’. ‘R’ followed by an object of reference expressed by a variable of individuals ( $x, y, z, \dots$ ), a proper name ( $m, n, o, \dots$ ), or an arbitrary name ( $a, b, c, \dots$ ) is read as “identifying reference is made to \_\_\_\_\_.”<sup>377</sup>

The property “ $\mathcal{R}$  is a class” is expressed by the axiom

A 1                     $\mathcal{C} \mathcal{R}$

P 1                     $\mathcal{R} \neq \emptyset$

I 2            P 1 is an “existence postulate” to the effect that  $\mathcal{R}$  is not empty. P 1 is complemented by P 2, P 6, P 7, and P 8.

P 1 *self-validates* in MoR as follows: By C 1,  $R\mathcal{R}$  is entailed. Assume the contrary, that  $\sim R\mathcal{R}$ . In this case,

---

<sup>377</sup> Identifying reference, here represented by ‘R’, is not a predicate, but, as we saw in {11.4}, may with more detail be expressed as a ternary relation between a person, a particular, and a space-time location (we recall from {11.4} that inclusion of a person is not presumed in every instance of identifying reference). For simplicity, this fuller expression in the form of a ternary relation is omitted here.

C 1 cannot be formulated. Therefore, the formulation of C 1 requires that  $\mathcal{R}$  itself be non-empty.

- C 2             $\mathcal{P}$
- I 3             $\mathcal{P}$  is the class of particulars. A member of  $\mathcal{P}$  is called ‘P’.
- A 2             $\subset \mathcal{P}$
- P 2             $\mathcal{P} \neq \emptyset$
- P 3             $\sim(\diamond a \wedge \sim \diamond a)$
- I 4            We leave open that to an object of reference complementary ascriptions may sometimes be made (see, e.g., {7.3.6, 27.8}). Nonetheless, the modal form of the law of non-contradiction, P 3, is accepted.
- Self-validates:* By appeal to the method of proof of self-validation itself,  $\diamond a \wedge \sim \diamond a$  is ruled out.
- T 1             $\diamond a \vee \sim \diamond a$
- I 5            This is a modal form of the law of excluded middle.
- Proof:* Assume  $\sim(\diamond a \vee \sim \diamond a)$ . Then  $\sim \diamond a \wedge \sim \sim \diamond a$  follows, which is ruled out by P 3.<sup>378</sup>
- P 4             $Ra \vdash \diamond Ra$       [Rule of possibility-introduction:  $\diamond I$ ]

---

<sup>378</sup> Some readers may wish to amplify this deduction by explicitly citing the interderivability postulate  $\sim \sim \diamond a \Vdash \diamond a$  and  $\diamond a \Vdash \sim \sim \diamond a$  in which ‘ $\Vdash$ ’ is the sign for metalogical entailment {11.4}. This postulate, the rule of double negation for possibility ( $\diamond DN$ ), expresses the concept of possibility fundamental to the metalogic of reference, i.e., what is possible relative to a system of reference is what is not prohibited ( $\sim \sim \diamond a$ ), and what is not prohibited is what is possible ( $\diamond a$ ). It is in this sense pertaining to constraints upon valid reference that we may understand the metalogical interderivability of what is not referentially forbidden and what is referentially permitted.

- I 6 From identifying reference to an arbitrarily selected object it follows that such reference is possible. (' $\vdash$ ' is the derivability sign, read as "therefore.")

*Self-validates:* Assume the falsifying case,  $Ra \wedge \sim\Diamond Ra$ .  $Ra$  provides a referential counter-example to  $\sim\Diamond Ra$ , therefore the falsifying case is dismissed. The formal metalogical paradigm for such referential counter-exemplification is given by:

- P 5  $\sim\Diamond Ra \Vdash Ra$  [Rule of referential counter-exemplification: RCE] (' $\Vdash$ ' is the sign for metatheory derivability, i.e., metalogical entailment; for discussion, see {11.4, 11.5}.)

- I 7 The denial that identifying reference to an arbitrarily selected object is possible metalogically entails identifying reference to it.

*Self-validates:* Identifying reference must be made to the arbitrarily selected object  $a$  in claiming that reference to  $a$  is impossible.

- T 2  $\sim\Diamond Ra \Vdash \Diamond Ra$

- I 8 The denial that reference is possible to an arbitrarily selected object metalogically entails that reference to that object is possible.

*Proof:* P 5, P 4.

- A 3  $\Diamond a \Vdash Pa$

- I 9 If an arbitrarily selected object is possible this metalogically entails that object is a possible object of reference.

*Self-validates:* The possible formulation of the left-hand side of A 3 metalogically entails that  $a$  is a particular.

T 3  $(x)\diamond Rx$

I 10 Identifying reference is possible to any object.

*Proof:* See I 11 below. There, we also obtain the associated result T 4:

T 4  $\sim(\exists x)\sim\diamond Rx$

I 11 There is no object to which identifying reference is impossible.

*Proof* of T 3 and T 4:

1	(1)	$(\exists x)\sim\diamond Rx$	A
1	(2)	$\sim(x)\diamond Rx$	Pred. calc.
3	(3)	$\sim\diamond Ra$	A
3	(4)	$Ra$	3 RCE
3	(5)	$\diamond Ra$	4, $\diamond I$
3	(6)	$\diamond Ra \wedge \sim\diamond Ra$	3, 5 $\wedge I$
	(7)	$\sim\sim\diamond Ra$	3, 6 RAA
	(8)	$\diamond Ra$	7, T 1, MTP
	(9)	$(x)\diamond Rx$	8 UI
1	(10)	$(x)\diamond Rx \wedge \sim(x)\diamond Rx$	2, 9 $\wedge I$
	(11)	$\sim(\exists x)\sim\diamond Rx$	1, 10 RAA

Lines (9) and (11) reach, respectively, the proof of T 3 and T 4.

T 5  $a \vdash \diamond a \vdash \diamond Ra$

*Proof:*  $\diamond I, A 3, D 1 (Pa =_{df} \diamond Ra)$ .

T 6  $\mathfrak{R} \neq \emptyset \Vdash (\exists x)(x = a \wedge Pa)$

I 12 There is at least one particular. That is, the class of identifying references is non-empty metalogically entails that there is at least one possible object of

reference. (Note that the denial of  $(\exists x)(x = a \wedge Pa)$  is equivalent to the rejection of P 1.)

*Proof:* P 1, P 4, D 1.

C 3  $\mathfrak{D}$

I 13  $\mathfrak{D}$  is the class of particulars such that  $\mathfrak{D} \subset \mathfrak{P}$  and such that  $\mathfrak{D}$  consists of properties and relations which (see the previous §5) can form the content of an ascription. A member of  $\mathfrak{D}$  is called 'D'.

A 4  $\mathcal{C}\mathfrak{D}$

P 6  $\mathfrak{D} \neq \emptyset$

Where D is a description, we define

D 3  $D =_{df} x: x \varepsilon \mathfrak{P} \wedge x \varepsilon \mathfrak{D}$

I 14 A description is a particular, a possible object of reference, consisting of a property or relation (§5).

T 7  $(\exists x)Dx$

I 15 There is at least one description.

*Proof:* T 6, P 1, P 6, D 3,  $\exists I$ .

T 8  $(\exists x)(Dx \supset Px)$

I 16 If something is a description then it is a particular, i.e., a possible object of reference.

*Proof:* From D 3, or T 7, T 5, D 1, CP,  $\exists I$ .

T 9  $\diamond Ra \Vdash (\exists x)(\exists y) (x \neq y \wedge Px \wedge Py)$

I 17 The possibility of identifying reference to an arbitrarily selected object of reference metalogically entails there are at least two particulars.

*Proof:* By T 6, P 1, T 3, and D 1 we know there is a possible object of reference, say  $Pa$ , and that thus  $\diamond Ra$ . (i) The last statement presupposes  $\diamond RPa$ , by D I or T 3. (ii)  $\diamond RPa$  is a particular distinguishable from  $Pa$ ; call it ' $Pb$ '. Statements (i) and (ii) correspondingly imply  $\diamond RPa$  and  $\diamond R Pb$ ; thus there are at least two particulars.

T 10  $\diamond Ra \Vdash (\exists x)(\exists y) (x \neq y \wedge Dx \wedge Py)$

I 18 The possibility of identifying reference to an arbitrarily selected object of reference metalogically entails that there are at least two objects of reference, a description and a particular. (I.e., at least one of the two particulars in I 17 is a description.)

Referring to I 17, possible reference to a particular ( $\diamond RP$ , which is a relation) itself comprises a description (D), hence T 10.

*Proof:* By C 3, P 6, C 2, and P 2, the content of what is ascribed ( $Dx$ ) in an identifying reference is distinguished from the object to which ascription is made ( $Py$ ). Identifying reference establishes that the content of what is ascribed applies to the object of ascription, as follows: Using ' $ADx$ ' to express 'ascription D is made of  $x$ ', we may then state

$$Ra \Vdash (\exists x)(\exists y) (Px \wedge Dy \wedge ADx \wedge x = y)$$

i.e., identifying reference to an arbitrarily selected object of reference metalogically entails that an ascription to that which can be the object of an ascription (i.e., a particular) establishes that what is ascribed and that that to which ascription is made are one and the same.

In this way, D 2 is translated into I 18.

C 4       $\mathfrak{F}$

I 19       $\mathfrak{F}$  is the class of identification frameworks. A member of  $\mathfrak{F}$  is called 'F'.

A 5       $\mathcal{C}\mathfrak{F}$

P 7       $\mathfrak{F} \neq \emptyset$

A 6       $\diamond Ra \Vdash F$

I 20      The possibility of identifying reference to an arbitrarily selected object metalogically entails an identification framework from the standpoint of which such reference is assured.

*Self-validates:* Assume the falsifying instance,  $\diamond Ra \wedge \sim F$ . Its assumption metalogically presupposes an F which permits  $R(\diamond Ra \wedge \sim F)$ . Therefore, A 6.<sup>379</sup>

T 11       $Ra \Vdash F$

I 21      Identifying reference to an object metalogically entails an identification framework.

*Proof:* P 4, A 6.

A 7       $F \supset \diamond Ra$

I 22      An identification framework provides a basis for possible identifying reference.

*Self-validates:* Assume the falsifying instance,  $F \wedge \sim \diamond Ra$ . By conjunction-elimination, T 2, P 3, conjunction-introduction, RAA, and basic prop. calc., A 7 is established. (Self-validation in MoR is

---

<sup>379</sup> It is immaterial here whether the latter F is identical to the F on the right-hand side of A 6.



considered to include indirect proof as a special case; theorems are proved without recourse to self-validation.)

T 12  $F \Vdash \Diamond Ra$  and  $\Diamond Ra \Vdash F$

I 23 An identification framework is metalogically equivalent to the possibility of identifying reference.

*Proof:* A 6, A 7, prop. calc.

T 13  $F \Vdash (\exists x)(\exists y) (x \neq y \wedge Px \wedge Py)$  and  $(\exists x)(\exists y) (x \neq y \wedge Px \wedge Py) \Vdash F$

I 24 An identification framework is metalogically entailed and is entailed by the existence of two particulars.

*Proof:* A 7, T 9, T 5, A 6.

A 8  $\Diamond R F$

I 25 Identifying reference is possible to an identification framework.

*Self-validates:* Similar to the self-validation of P 1, applying P 4 ( $\Diamond I$ ).

T 14  $\Diamond R F \Vdash F'$

I 26 The possibility of identifying reference to an identification framework itself metalogically entails an identification framework. (No assumption is made here that F and F' must be distinct frameworks.)

*Proof:* A 6 (S). (Substitution)

D 4  $\Diamond^m R \stackrel{\text{df}}{=} \Diamond R (\Diamond R \dots \Diamond Ra)$

T 15  $\Diamond Ra \Vdash \Diamond^m R$

I 27 The possibility of identifying reference to an arbitrarily selected object metalogically entails the possibility of iterative references.

*Proof:* By  $m$  applications of T 5 (S).

D 5  $F^m =_{df}$  the identification framework metalogically entailed by  $\diamond^m R$

T 16  $\diamond Ra \Vdash F^m$

I 28 The possibility of identifying reference metalogically entails an identification framework from the standpoint of which iterative references are possible.

*Proof:* T 15, D 5.

T 17  $Pa \supset \diamond Ra$

I 29 If something is a particular, it is possible to refer to it.

*Proof:* By T 3.

T 18  $\diamond a \Vdash F^m$

I 30 The possibility of a particular metalogically entails an identification framework from the standpoint of which iterative references are possible.

*Proof:* A 3, T 17, T 16.

T 19  $a \Vdash F^m$

I 31 An arbitrarily selected object of reference metalogically entails an identification framework from the standpoint of which iterative references are possible.

*Proof:* T 5, T 16.

## §7. De-projection and the ontology of MoR: Informal comments

In previous chapters of this book, I have tried to give the reader a detailed understanding of the scope, objective, and methodology of the metalogic of reference. A route has been described which has led from an initial interest in the approach to conceptual analysis proposed by transcendental philosophy, to an understanding of the concept of possibility in terms of metalogical conditions of valid reference. The concepts basic to a study of such conditions have been discussed in detail and in the preceding section an important group of these has been given formalized expression for the purpose of clarifying certain of the important relations between them.

The valid employment of these concepts is, however, subject to various constraints. A description of these constraints serves to represent the theoretical limits of reference and meaning, which in this work have been called '*horizons*'. As we have seen, in attempting to transgress beyond such horizons, any theory (or concept, claim, or position) lapses into meaninglessness. Since the metalogic of reference is itself a theory, moreover a theory on the level of maximum theoretical generality, a description of these essential theoretical limitations serves to define the sense in which its results are self-applicable and correspondingly subject to the same fundamental framework-relative limitations.

The objective of the metalogic of reference is achieved by understanding the structure of valid reference in relation to ways in which reference undermines itself on the level of possibility and becomes meaningless. The self-undermining of reference on the level of possibility then defines a task, that of providing a morphology of metalogical projections, a *pathology* of the forms that such dysfunctional reference may take; to elaborate such a pathology has been the goal of this *Critique of Impure Reason*. As we've seen in earlier chapters, it has been possible to find preventative, revisionary, and eliminative measures to avoid such projective departures from meaningfulness. In the process, the methodology of the metalogic of reference makes clear those sets of injunctions with which valid reference must *necessarily* comply, where such necessity is understood as the metalogical necessity conditioning possible reference. When the necessity of a set of referential injunctions is made evident, limiting conditions of possible reference and of meaning are made explicit.

In short, by determining the most fundamental constraints upon valid reference, the domain of possible sense is defined *negatively*. The metalogic of reference makes clear what is *forbidden*: all else is permitted. There is an analogy here with the definition of games: Were one to attempt to give the in-

structions for playing any very complicated game by listing all of the moves that are *permitted*, the set of instructions can become so long the game could not be played.<sup>380</sup> As a result, we often are forced to describe a game in terms of the *constraints* placed upon possible moves. Were we think of the Ten Commandments as defining a “moral game,” it is significant that, as noted in {7.5}, eight out of the 10 prohibitions take the form of ‘Thou shalt *not*...’.

To the end of defining the domain of possible reference and meaning, the method of de-projection, as we’ve seen, makes it possible to detect and enforce the fundamental constraints upon valid reference relative to a given identification system. As shown in previous chapters, when projections are identified and eliminated by mean of de-projection, metalogical analyses introduce no supplementary content or assumptions; as a result, the formal structure of de-projection is tautologous: In making explicit the metalogical constraints with which what is described must comply, a point is reached where by self-validation it is found that a concept (or claim, position, or theory) tautologically entails its own referential preconditions. It is precisely because de-projection is empty of content that it can authorize a transition from an initial projective claim to a de-projected revision, in the process often preserving the referential framework presupposed by that claim and without risking the introduction of error.

As we have seen in {18}, the ontology of the metalogic of reference is an “open ontology.” In contrast, a “closed ontology” has two characteristics: Such an ontology (i) accepts as “legitimately ‘existing’ entities”—that is, entities that in some specified framework-relative sense are said to *be* objects of reference—only those objects of reference that can be grouped together in some comparatively well-defined, or possibly indeterminate, set *O*, and (ii) rejects the claim that there *are* “legitimately ‘existing’ objects” not contained in *O*, independently of the reference frame presupposed by reference to *O*. Let *a* be such an object not a member of *O*. To deny that *a* is an object of reference is of course projective since reference must be made to *a* in order to reject *a* as a legitimate object of reference. The only justification for *closing* one’s ontology—for limiting, in other words, the kinds of objects of reference that *are* to be accepted—is that one refuses to admit the legitimacy of objects of reference beyond those of a certain sort. Often—as we have seen, for example, in the case of Quine ({3.2})—this refusal is no more than “a matter of taste” rather than a theoretically compelling result.

Although the ontology of MoR is open or ontologically neutral, it must accept certain existential-referential commitments. That this is the case is seen from the now-familiar fact that any attempt to deny that objects of a certain

---

<sup>380</sup> (Not, at least, by an unaided human being.)

kind *are*—in whatever sense a given framework of reference sanctions—leads to metalogical self-referential inconsistency (again, see {18}). Although the realist assertion that certain classes of objects possess the autonomy of total framework-independence, and the idealist assertion that certain classes of objects are inherently framework-dependent, are both metalogically self-undermining and hence without possible meaning ({21}), objects in general, as they are and can be understood reflectively, are inescapably functionally interdependent in relation to reference frames that permit reference to them.

The general ontological conclusion that MoR draws from the family of projections considered in this study is this: Objects of possible reference *are*—in principle and therefore in fact—a function of those identification frameworks which provide the basis for identifying reference to them. MoR recognizes that modes of referring functionally determine domains of particulars, that any object possessing a specifiable identity is relative to identification systems which in principle provide the means to refer to it, and that any concept, claim, position, or theory is without meaning if it purports to refer to certain objects while undercutting the basis necessary to permit possible reference to them.

### §8. Formalization continued

The transition from descriptions of preconditions of reference and of non-projective concepts to their formal expression is normally accomplished by turning attention from *concepts* to a corresponding set of formalized *propositions* (henceforth, or statements). The formalized representation of the concept of projection therefore requires that we consider appropriate ways of handling these formalized propositions with which a desired interpretation can be associated.

A decision that has always to be made before one seeks to give formal expression to a theory concerns the nature of the propositions to be represented in the formalism. Specifically, there is a need to determine whether the propositions in question are all of them such that the values “truth” and “falsity” exhaust the values such propositions can, in the intended interpretation, assume. If this turns out to be the case, then a two-valued logic can be employed in the formalization. Accordingly, the principle of bivalence (traditionally, the law of excluded middle) would be asserted to hold in the formalization, and a given formalized statement would assume either of the two values, but no others.

As we recall from {11.4}, in order to take metalogically self-undermining statements into account in a formal expression of the metalogic of reference,

we have need of at least a three-valued logic, one whose significant range comprises at least the values truth (T or 1) and falsity (F or 0), and with ‘ $\mu$ ’ representing the value of projective assertions: ‘ $\mu$ ’ represents the value “projective meaningfulness” which lies outside the range of significant values. The statements or propositions represented in MoR would then be at least three-valued. —However, as we saw also in {11.4}, *from a metatheoretical standpoint* in terms of which it is possible to evaluate the referential consistency of individual statements or propositions, *bivalence is accepted*: From that standpoint, a proposition either is projective or it is not; if not, it may have a value in the significant range, truth or falsity.

Furthermore, it can be challenging in a formalized language to deal adequately with the conceptual issues that relate to the problem of putative meaning. As remarked in {12.1}, we face a task that is potentially elusive when we seek to represent in symbolic form the conceptual misunderstanding—i.e., an erroneous and deceptive *understanding*—that a proposition is devoid of meaning which was initially thought to be meaningful. We recall that the formalization given in {11.4} approached the problem of putative meaning in terms of *putatively referring propositions* (or sentences).

To view the problem of putative meaning from a somewhat different perspective, in this Supplement we consider an alternative, complementary approach by instead expressing the issue of putative meaning in terms of *putative objects of reference*. In {11.4}, a proposition is projectively self-undermining when it involves a *putative reference* that undercuts its own possibility. Alternatively here we may consider a proposition to involve projection when it makes claims about “putative objects of reference”—that is, claims *seemingly* about objects of reference, but “objects” to which reference is impossible—“objects of reference” only in an illusory sense that is without meaning.

We may express this formally by using the variable ‘ $o$ ’ to stand for a blank space-holder, a blank that *would be filled* were an object to have an identity such that it could in principle be identified.<sup>381</sup> A projection then takes the same form as we saw in {11.4}:

$$\vec{p} \equiv A(o) \vee \sim A(o). \wedge \sim \diamond R_o \quad (\text{condition 1})$$

which states that a projection is involved iff an ascription A or an ascription  $\sim A$  is made of a putative object of reference  $o$ , while reference to  $o$  is impos-

---

<sup>381</sup> It is important to recognize that such a blank space-holder is of course of an altogether different sort from the name of an arbitrarily selected object, as T 19, for example, should make evident.

sible. Under these conditions, the place-holder  $o$  cannot in principle be filled by an object. In this way we express in symbolic form the delusion-inducing sense in which projections *purport* to have objects of reference in view.

Unlike the place-holder  $o$ , in what follows we continue to let ‘ $a$ ’ represent an arbitrary name of an individual with some identity, and read  $Ra$  as “identifying reference is made to  $a$ .” We then add the following to the formalized expression of MoR given in §6:

C 5             $\mathcal{S}$

I 32             $\mathcal{S}$  is the class of three-valued propositions (or statements) which may assume the values T, F, or  $\mu$ . An individual proposition or statement belonging to  $\mathcal{S}$  is called ‘ $p$ ’.

A 9             $\mathcal{C}\mathcal{S}$

P 8             $\mathcal{S} \neq \emptyset$

I 33            P 8 is an “existence postulate” to the effect that  $\mathcal{S}$  is not empty.

P 8 *self-validates* in MoR as follows: P 8 itself qualifies as a  $p_i$  belonging to  $\mathcal{S}$ .<sup>382</sup>

A 10             $(x)(x = p \supset p \in \mathcal{P})$

I 34            Any proposition  $p$  is a member of the class of particulars—i.e., an individual proposition is itself a possible object of reference,  $P_{p_i}$ .

T 20             $p_i \Vdash F_i$

---

<sup>382</sup> In this connection, we note that the three-valued system of Bochvar (see {11.4}) does not require a theory of types, while Moh Shaw-kwei’s efforts seem to indicate that Russell’s paradox, in the form  $\lambda x.x \notin x$ , cannot be generated in a three-valued system of this kind. See Church (1939-40) and Moh (1954, p. 37).

I 35 An individual proposition  $p_i$  metalogically entails an appropriate identification framework which provides a basis for  $p_i$ 's capacity to refer identifyingly. An appropriate identification framework satisfies T 11 and T 12.

*Proof:* C 2, A 10, T 12.

D 6  $\bar{p}^\rightarrow$  for any  $p$  such that  $p \equiv A(o) \vee \sim A(o), \wedge \sim \diamond Ro$

I 36 An individual proposition is projective as stated earlier under condition (1).

D 7 T for a  $p_i$ , where ' $p_i$ ' iff  $p_i$ . (Tarski's definition)

D 8 F for a  $p_i$  where ' $p_i$ ' iff  $\sim p_i$ .

D 9  $\mu$  for a  $p_i$ , where ' $p_i$ ' iff  $\bar{p}_i$ .

T 21  $\bar{p}_i \Vdash \sim F_i$

I 37 A projective proposition metalogically entails a denial of the identification framework it requires for its possible reference. Such a metatheoretical denial amounts to the assertion that there does not exist an  $F_i$  which is required in order for  $\bar{p}_i$  possibly to refer.

*Proof:* D 6, T 12.

T 22  $\bar{p}^\rightarrow \supset \diamond Ro \wedge \sim \diamond Ro$

I 38 A projective proposition exhibits a modal contradiction of the type prohibited by P 3: specifically, a projective proposition is metalogically self-undermining.

*Proof:* Since by D 6 an individual  $\bar{p}_i$  claims to make a true or false ascription to an (alleged) object of



reference  $o$ ,  $\bar{p}_i \supset Ro$  follows, and by P 4,  $\bar{p}_i \supset \Diamond Ro$ . But, in addition to its ascription claim, by D 6,  $\bar{p}_i \supset \sim\Diamond Ro$ . Therefore, T 22.

By P 3, any putatively referring  $\bar{p}^\rightarrow$  must be rejected since under the conditions of its alleged truth-functional ascription A or not-A, it presupposes an alleged object of ascription  $o$ , and hence  $Ro$ , but it rules out  $\Diamond Ro$ ;  $o$  is therefore a blank space-holder with which no meaning can be associated.

Under these conditions, by P 3 and the results reached in {11.4} and {17}, there is no other rational alternative but to reject (and not simply deny) any  $\bar{p}^\rightarrow$ .

Using the metalogical rejection symbol ' $\neg\|\$ ' defined in {11.4}, we therefore conclude  $\neg\|\bar{p}^\rightarrow$ .

**§9. Schematic summary of formalization  
found in this Supplement**

Supplement §6:

C 1	ℜ	D 1	Particular	A 1	℄ℜ
C 2	℘	D 2	Identifying reference	A 2	℄℘
C 3	℔	D 3	Description	A 3	℔a ⊢ Pa
C 4	ℑ	D 4	℔ <sup>m</sup> R	A 4	℄℔
		D 5	F <sup>m</sup>	A 5	℄ℑ
				A 6	℔Ra ⊢ F
				A 7	F ⊃ ℔Ra
				A 8	℔RF

P 1    ℜ ≠ ∅

P 2    ℘ ≠ ∅

P 3     $\sim(\diamond a \wedge \sim\diamond a)$

P 4    Ra ⊢ ℔Ra

P 5     $\sim\diamond Ra \vdash Ra$

P 6    ℔ ≠ ∅

P 7    ℑ ≠ ∅

T 1     $\diamond a \vee \sim\diamond a$

T 2     $\sim\diamond Ra \vdash \diamond Ra$

T 3    (x)℔Rx

T 4     $\sim(\exists \diamond x)\sim\diamond R \diamond x$

T 5    a ⊢ ℔a ⊢ ℔Ra

T 6     $\mathfrak{R} \neq \emptyset \vdash (\exists x)(x = a \wedge Pa)$

T 7    (∃x)Dx

T 8    (∃x)(Dx ⊃ Px)

T 9    ℔Ra ⊢ (∃x)(∃y)(x ≠ y ∧ Px ∧ Py)

T 10    ℔Ra ⊢ (∃x)(∃y)(x ≠ y ∧ Dx ∧ Py)

T 11    Ra ⊢ F

T 12    F ⊢ ℔Ra and ℔Ra ⊢ F

T 13    F ⊢ (∃x)(∃y)(x ≠ y ∧ Px ∧ Py) and (∃x)(∃y)(x ≠ y ∧ Px ∧ Py) ⊢ F

T 14    ℔RF ⊢ F'

T 15    ℔Ra ⊢ ℔<sup>m</sup>R

- T 16  $\diamond Ra \Vdash F^m$   
 T 17  $Pa \supset \diamond Ra$   
 T 18  $\diamond a \Vdash F^m$   
 T 19  $a \Vdash F^m$

Supplement §8:

- |     |               |     |                       |      |  |
|-----|---------------|-----|-----------------------|------|--|
| C 5 | $\mathcal{S}$ | D 6 | $\vec{p}$ Projection  | A 9  | $\mathcal{C}\mathcal{S}$                     |
|     |               | D 7 | T Truth               | A 10 | $(x)(x=p \supset p \varepsilon \mathcal{P})$ |
|     |               | D 8 | F Falsity             |      |  |
|     |               | D 9 | $\mu$ Meaninglessness |      |  |
- P 8  $\mathcal{S} \neq \emptyset$
- T 20  $p_i \Vdash F_i$   
 T 21  $\vec{p}_i \Vdash \sim F_i$   
 T 22  $\vec{p}_i \supset \diamond Ro \wedge \sim \diamond Ro$

### §10. A concluding reminder

In MoR, the axioms, rules of derivation, and deductions made from them are formalized expressions which we may understand, when they are employed together, as schematisms of heuristic inference. It is presumed—indeed in light of the plausible consequences of §§2–3 in this Supplement, it would be *required*—that reflective, analytical intelligence of a certain kind needs to be applied in any deductive extension and use of the formalized heuristics of MoR. If the plausible conclusions are correct which were reached in §§2–3 concerning the possible in-principle deductive or “non-deductive” non-formalizability of the metalogic of reference (from which it would follow that any attempted formalization would *necessarily* be inadequate), then the possibility is left open that “*unintelligent*” or blindly mechanical, rule-driven derivations from the formalized heuristics may potentially lead to paradoxical and/or inconsistent inferences. This should neither surprise nor perturb us, although clearly this would be a disappointment to anyone who wishes for the advantages of a system that is completely formalizable and equipped with an

effective decision procedure, but it would be a disappointment that follows from the very limitations of formalization that we have considered. Nevertheless, it is conceptually bracing as well as healthy for us to remember Minsky, who did not believe that consistency is always necessary or even desirable: “What is important is how one handles paradox or conflict, how one learns from mistakes, how one turns aside from suspected inconsistencies...” (Minsky, 1974, p. 76).

Given that the metalogic of reference deals with a widespread conceptually misleading, delusion-inducing variety of self-undermining reference, there is an indispensable need to recognize the essential role of what I have called ‘*epistemological intelligence*’ both in understanding and in applying the formalized principles sketched in §§6 and 8 of this Supplement. A detailed description of many of the principal defining characteristics of epistemological intelligence will be found in Appendix II.

# APPENDIX I

## The Concept of Horizon in the Work of Other Philosophers

[I]n order to be able to set a limit to thought, we should have to find both sides of the limit thinkable (i.e. we should have to be able to think what cannot be thought).

– Ludwig Wittgenstein (1961/1921, p. 3)

It is not ... clear how we can make out from within the limits to our own imaginative and constructive intellectual powers.

– Humphrey Palmer (1985, p. 144)

The concept of horizon is a very general, flexible, accommodating idea, one which has been put to many uses both in and beyond philosophy. As we have seen in {14}, it is frequently used in its ordinary geographical or physical sense to designate the limit of the Earth's surface visible from a certain location. But the idea has been employed in many other ways, for example, to describe the compass—the range of inclusion—of a thought, of a point of view, and even of an action. The concept of horizon is also used in astronomy (e.g., to refer to the plane of the great circle of the celestial sphere that passes through the center of the Earth, parallel to the sensible horizon at a given location). Horizons appear in archeology, to distinguish levels at which artifacts from a given period are found, and also in anatomy and zoology—in connection with the retinal horizon, or that of the diaphragm or of the teeth. In geology, horizons are levels of stratification that are thought once to have been continuous horizontal strata, or they may refer to a distinguishable level at which certain fossils are found. The concept of horizon is very versatile and lends itself to loose as well as more precise applications.

And so it is no surprise that in the writings of philosophers, the concept of horizon and words for it in a variety of languages have been put to many

different and highly varied uses. It would require a book in itself to describe the many meanings that philosophers have associated with the idea. Here, I want only to mention, briefly and summarily, two groups of philosophers who have employed the concept of horizon in ways relevant or at least related to the concept of horizon studied in this book.

### §1. The phenomenological horizon

German philosopher Edmund Husserl (1859–1938) used the concept of horizon in a number of ways, as have other phenomenologists. Like so much in phenomenology, there is a good deal of vagueness that permeates its terminology, and an equally good deal of effort (and space) would be needed to disentangle the different meanings to which the term ‘horizon’ has been assigned by phenomenologists. Here, I will mention only two principal ways in which Husserl employed the concept.

Husserl was significantly influenced by the work of William James, and in connection with Husserl’s concept of horizon, he was influenced by James’ “doctrine of fringes.”<sup>383</sup> In James’ words: “Our fields of experience have no more definite boundaries than have our fields of view. Both are fringed forever by a more that continuously develops, and that continuously supersedes them as life proceeds” (James, 1958/1904, p. 71). This implicitly “expandable” conception of horizon we then find in Husserl work, but it is put to different and more specific purposes.

In his phenomenology of time consciousness, Husserl applied the concept of horizon to describe the way in which experience of the present moment has both a “retentive” and a “protentive” constitution: As experience continues, the present is surrounded by “horizons” (or “halos”) of retained memories of the past, and by anticipations of the future.<sup>384</sup> Another application of the concept of horizon is found in Husserl’s description of perceived objects: As Bachelard (1957, p. 251) expressed this: “[P]erception of such an object possesses a horizon that embraces other perceptual possibilities. And these possibilities are implied by the sense of the perceived object.”<sup>385</sup>

---

<sup>383</sup> For a discussion of this influence, see Spiegelberg (1965, I, pp. 111-117).

<sup>384</sup> See Husserl (1928, for example §§24, 29).

<sup>385</sup> Author’s translation. “[L]a perception... [d’un tel] objet possède un horizon qui embrasse d’autres possibilités perceptives. Et ces possibilités sont impliqués dans le sens de l’objet perçu.”

Bartlett (1970) examines, from a phenomenological perspective, the underlying constitutive structure (there called ‘logic of structure’) of particular objects of reference (Section 1.5,

Both of these applications of the concept of horizon take advantage of its ordinary physical meaning as a potentially expandable or continuable scope of vision, and then extend that sense as an aid in understanding how the human consciousness of the passage of time is made possible by the way in which present experience comes, so to speak, with built-in “fringes” or “halos” or “horizons” that make it possible to retain memories of the past and anticipations of the future. Both kinds of “horizon” in Husserl contain the potential to enlarge or expand as experience proceeds.

A more general, inclusive phenomenological application of the concept of horizon was expressed by Kuhn (1940, pp. 107-108):

(1) Horizon is the ultimate circumference within which all things, real and imaginable, are found to appear. To explore the horizon means to move away from the ordinary foci of attention with a view to integrating the things at hand in a broader and ever broader context. The idea of horizon stands for the progressive drive inherent in experience. (2) While limiting the totality of given things, the horizon also frames it. The frame of a picture, though forming no part of it, helps to constitute its wholeness. Similarly, the horizon determines that which it frames.... (3) By its very nature every horizon is “open.” As we move from the center toward the circumference fresh horizons open up. We are constantly invited to transcend the boundary of our field of vision.... Thus the notion of horizon points to a basis of experience outside experience. It stands for the impetus of self-transcendence with which experience is animated.

This notion that a horizon “points to a basis of experience outside of experience” we also found expressed by Kant in his “feeling of being forced” to go beyond the limits of possible experience (see {14}).

Physicist-philosopher Patrick Heelan considered a certain phenomenological concept of horizon to be useful in providing a realist understanding of quantum theory. In that context, it was for Heelan a broad and general concept. As he described it:

The attempt to elucidate and criticize the kind of cognitive intentionality-structure implicit (and sometimes—but rarely—

---

“Valences”) and of temporal experience (Section 2.1, “The Structure of Phenomenological Space-time”).

explicit) in the form of life which characterizes a given type of empirical scientific investigation is called an analysis of the horizon of the science.” (Heelan, 1967, p. 379).

From these brief examples, the reader will quickly see that no single unitary concept of horizon is to be found among its phenomenological applications.

## §2. The horizon of transcendental Thomists

Canadian Jesuit philosopher Bernard Lonergan (1904–1984) began with the ordinary geographical-physical concept of horizon—“a maximum field of vision from a determinate standpoint” (Lonergan, 1988/1967, p. 198)—but quickly extended the idea to mean what is commonly considered to be a philosopher’s total conceptual framework. Much like Carnap’s concept of an entire linguistic system (discussed in {13.4}), Lonergan’s horizon can be resistant to being called into question from the standpoint of external frameworks or conceptual systems. Here was Lonergan’s view:

In a generalized sense, a horizon is specified by two poles, one objective and the other subjective, with each pole conditioning the other.... [H]orizon is prior to the meaning of statements: every statement made by a realist denotes an object in a realist’s world; every statement made by an idealist denotes an object in an idealist’s world; the two sets of objects are disparate; and neither of the two sets of statements can prove the horizon within which each set has its meaning, simply because the statements can have their meaning only by presupposing their proper horizon. (Lonergan, 1988/1967, pp. 198-199)

Lonergan also described the idea of horizon in more phenomenological terms: “...each of us lives in a real world of his own. Its contents are determined by his *Sorge*, by his interests and concerns, by the orientation of his living, by the unconscious horizon that blocks from his view the rest of reality” (p. 148). In both senses, this more experiential one and the earlier notion that a horizon comprises a philosopher’s total conceptual framework, we notice that the concept of horizon is limiting: in the first case, such a horizon “blocks” an individual’s “view of the rest of reality”; in the second case, the “horizon” of a philosopher’s position is hermetically sealed within its own



system of presuppositions.<sup>386</sup>

Lonergan related his concept of horizon to its use by Austrian Jesuit philosopher Emerich Coreth (1919-2006). Coreth, like Lonergan, contributed to the development of what has come to be known as transcendental Thomism, among whose other proponents were Gaston Isaye and Joseph Maréchal, whose work was discussed in {6}. Coreth was interested in applying transcendental argumentation to bring to light contradictions between judgments and the “implicit metaphysics” they presuppose. He advocated strongly for the philosophical usefulness of the transcendental approach:

[T]ranscendental method, as we understand it, is not only the fundamental method that is demanded by the nature of metaphysics as basic science; it is also, one might venture to say, the integral method that takes over all other methods which, standing in isolation from one another, are insufficient, takes them over and, while respecting their legitimate concerns, sublates them into a higher unity” (Coreth, 1961, p. 88).

It was in this context that Coreth applied his idea of horizon. Coreth’s concept of horizon, as Lonergan described it concisely, “is total, for beyond being there is nothing. It is basic, for a total horizon is basic; it cannot be transcended, gone beyond, and so it cannot be revised” (Lonergan, 1988/1967, p. 200). Depending upon one’s philosophical sympathies, this says a lot or very little. But it is characteristically phenomenological.

---

<sup>386</sup> Yet another use to which Lonergan put the concept of horizon is discussed in Rehg (1989).

## APPENDIX II

### Epistemological Intelligence

This appendix continues research described in earlier publications by the author relating to the psychology of philosophers.<sup>387</sup> Here, I identify a set of important skills essential to what I will call ‘*epistemological intelligence*’. These skills are examined in relation to a group of psychological characteristics shared by many philosophers and by many students attracted to the study of philosophy. The twofold purpose of this appendix is to recognize epistemological intelligence as a distinguishable variety of human intelligence, one that is especially of value to philosophers, and to understand the challenges posed by the psychological profile of many philosophers and its students that can impede the development and cultivation of the skills associated with this variety of intelligence.

#### §1. Two approaches to the study of epistemology

The subject of epistemological intelligence, as I hope to make evident, is important, seldom discussed, and deserves study in its own right; it is also a form of intelligence relied upon by the approach to the metalogical analysis of reference with which this book has been concerned.

On the one hand, there is a form of human thinking that is essentially epistemological. On the other, there is the subject of study, epistemology, the product of epistemological thought. The former is not epistemology, but it is a precondition for epistemology’s successful realization. The latter is that realization, sometimes successful, sometimes not. The two belong to very different species. It is the first, epistemological thinking and the skills that make it possible, that will concern us here.

The most common definition of epistemology is “the study of the limits

---

<sup>387</sup> Bartlett (1978-79, 1986a, 1986b, 1989).

and conditions of knowledge.” There are two very distinct ways of studying such “limits” and “conditions.” In most universities, epistemology is taught as a *content*-based subject-matter: In this approach, individual epistemological theories that have been formulated by historically famous philosophers are described and critiqued. Relatively seldom is epistemology taught as a *skill*-based discipline—as, literally, a reflective mental discipline, understood in terms of a specific set of well-defined skills, whose purpose is to enable the identification and clarification of those “limits and conditions of knowledge” by employing the cognitive resources of the individual student.

If the objective of a class in epistemology were to focus on the particular cognitive skills that are epistemologically important, the individual student would ideally first learn to distinguish and to separate, without undue self-conscious effort, his or her own mere beliefs, on the one hand, from states of solid and reliable knowledge, on the other. But not only this, and again ideally, the student would come to recognize the need to revise his or her fundamental ways of understanding and conceptualizing the world, and undertake to do this by means of a two-step process that discards unanalyzed, previously accepted, baseless beliefs, and then replaces them with a set of reflectively analyzed and justifiable claims.

This approach to epistemology is not only less familiar to most professional philosophers, for many it is unknown. The notion that there is a specific set of epistemologically valuable “well-defined skills” is likely to be questioned, and by some philosophers also opposed. A skill-based approach to epistemology in the sense just described itself comprises a philosophical position that for some stands in need of philosophical justification.

Recognizing the legitimacy of a skeptical response like this, I need to make my modest intentions in this appendix clear. There are two relevant and important issues: On the one hand, we may wish to have an adequate philosophical justification for specifying the particular skills that I will identify. This is a justification that the teacher of the specified skills must be able to provide. On the other hand, we may also be interested in the degree to which philosophy students and professional philosophers may encounter difficulties in developing and applying those skills.

My objectives here are focused and therefore limited: As is common in mathematics, I propose to ignore the first issue by reducing it to a problem previously solved: that is, to refer readers to certain of my previous publications if they wish to know the philosophical justification for the approach that is in view.<sup>388</sup> Here, my interest is instead in the second issue: variability in the *epistemological abilities* of individual students and of professional philoso-

---

<sup>388</sup> See Bartlett (1970, 1975, 1976, 1982, 1983a, 1983b, 2005, 2011, 2015, 2016).

phers, a subject that will lead us to examine the nature of *epistemological intelligence*.

In much of my university teaching and in many of my publications, I have focused on epistemology in this skill-based sense. In the process, it has become clear to me that comparatively few people are able to develop the relevant skills in a proficient and lasting way.

This was initially a depressing realization—certainly an instance in which facts fail to satisfy what one would prefer. But once the fact of individual cognitive differences made itself evident in a skill-based epistemological context, it was then only a small step to ask, Why is this the case? and What does this reveal about the workings of the individual philosophizing mind?

## §2. A set of epistemological skills

The following are among the main reflective thinking skills I have sought to communicate to students and to readers of some of my published work; certain of these skills qualify, in my view, as epistemologically “special” because they are largely unique to epistemological study and are relatively unknown and unused outside of epistemology. In the list below, I am not concerned to provide an exhaustive enumeration of all relevant skills, nor to establish the independence of each skill from the others:

- (1) a commitment to logical coherence and alertness to inconsistency
- (2) a commitment and a sensitivity to identify beliefs which are baseless, that is, which have no empirical or logical justification
- (3) a commitment and a will to eliminate, to the extent that this is possible, baseless beliefs from one’s own thinking
- (4) a commitment to seek revisionary replacements, again to the extent that this is conceptually possible, for beliefs recognized as baseless, and a determination to accomplish this
- (5) an ability to recognize that all claims to knowledge are inextricably framework-relative, that is, cannot be asserted when disassociated from the conceptual framework(s) that must be accepted in order for it to be possible to make those claims
- (6) a heightened awareness of the ways in which some scientific and many commonly accepted and widely used concepts *trespass* beyond the frameworks they presuppose, and by doing this, become *self-destructively incoherent, i.e., metalogically projective*.

- (7) a “*mental dynamic*” of a certain sort, one that has been developed by the individual, perhaps through training or perhaps simply because the person finds it natural and desirable to invest credence only in rationally justifiable ways of understanding and rationally justifiable claims that can be made on this basis
- (8) the establishment of what I have previously<sup>389</sup> called a ‘*rational bridge*’, which connects, on one side, the preceding reflectively, self-consciously justified and formulated understanding with, on the other side, both the individual’s own predispositions to think and behave in certain ways, and actual cognitive and behavioral conduct consistent with those predispositions

These eight basic skills are evidently of different kinds: (1)–(4) and (8) each explicitly combines distinct psychological and cognitive commitments with specific skills. (1) combines a certain variety of commitment with an individual’s *alertness* to inconsistency. Alertness can be considered a species of ability or of skill, which, like all skills, is susceptible to variation among individual people. (2) also combines a certain focus of commitment with a form of *sensitivity* to baseless beliefs, again a variable human ability. (3) combines a type of commitment with a strong degree of *determination*, also a kind of ability, to eliminate baseless beliefs from one’s thinking. (4), too, bonds commitment with an individual’s degree of *resolve* to find revisionary replacements. (5) refers explicitly to a specific ability, the ability to *recognize* framework relativity, again a skill that varies among individuals. (6) refers to a *heightened awareness* of instances of “framework transgression,” also a skill-based proficiency. (7), as well, points to an individual’s ability to establish within his or her mind a *conceptual habit* or pattern of thinking that leads the person to *believe* only when there is a justifiable basis for such belief. (8) is, I have observed, in the nature of a “second-order skill” which is the consequence, at least for some individuals, of the acquisition, solidification, and strengthening of the preceding seven “first-order skills.”

Some of these skills, abilities, or proficiencies are clearly more specialized in terms of their application within epistemology than are the others. Many people, especially scientists, develop skills of the first three kinds: They are committed to logical coherence, and are alert to inconsistencies in their thought and research results; they have learned to develop a sensitivity to empirically baseless or mathematically unfounded claims; and they are

---

<sup>389</sup> Bartlett (1969-70, 2016) and in the present book {1.1, 17.3, 29.4}.

committed to eliminating baseless beliefs and claims that are made as a result of them.

The fourth skill tends to be found in more creative individuals, and less among conventional thinkers and researchers who do not push the boundaries of conventionally accepted frameworks and do not attempt to initiate revisionary or revolutionary approaches.

The fifth skill, that of recognizing the framework relativity of claims, is today, I suggest, found primarily among physicists whose work focuses on relativity and quantum theory; in both of these areas of study, framework relativity is not only recognized, but, as we have seen in {26–28}, is firmly embedded in the very fabric of special and general relativity and of quantum theory. But in its most general, theoretically abstract form, the skill of recognizing the framework relativity of knowledge claims is essentially an epistemological skill.

The sixth skill—a heightened awareness of ways in which we tend to trespass beyond the boundaries of the frameworks we presuppose—is a skill I have tried in my university teaching and in various publications explicitly to communicate and to encourage others to make part of their thinking. It is a skill that, in my experience in teaching and writing, is difficult to impart and to encourage in individuals, for reasons I will get to shortly.

The seventh skill—to limit one’s credence to the justifiable—is a commonly advocated skill, whether among professional philosophers, or skeptics, or scientists and mathematicians. But when this skill is tied to the previous ability of recognizing framework relativity, it is not often cultivated by many people; indeed, I have found, by very few.

The eighth and last skill—really the establishment of a direct and reliable connection between an individual’s rational skills and *understanding*, and his or her cognitive and behavioral *conduct*—is also, again in my observation, rather rarely found among people.<sup>390</sup>

The set of eight skills that I’ve identified, *when they are combined together by a single mind*, some of them common to professionals in various fields and some more epistemologically centered, constitute what I shall call a ‘minimal set’ of *epistemological abilities*.

---

<sup>390</sup> On the surface, this appears to express a mere subjective opinion. It does not. There is a strong evidence that supports it. See Bartlett (2011, “The Distribution of Mental Health,” pp. 273-276), which charts the distribution of both positive and negative mental health. The higher, positive end of the spectrum includes those comparatively rare individuals in whom a rational bridge has formed that links their rationality with consistent behavior in the sense described in the text.

### §3. From a defined set of skills to the recognition of a new variety of intelligence

Investigators of human intelligence have taken as fundamental the close association of specific abilities or skills with corresponding distinct varieties of intelligence. To give a few historical examples: In 1895, Alfred Binet and Victor Henri were critical of then-existing intelligence tests due to the tests' tendency to oversimplify; Binet and Henri sought to persuade other intelligence researchers to develop a diversity of tests to measure distinguishable human mental abilities, including, for example, tests of imagination, attentiveness, memory, richness of mental imagery, verbal and mechanical comprehension, aesthetic appreciation, and even the capacity to sustain muscular effort and moral sensibility (Binet and Henri, 1895). They proposed ways of testing these various categories of 'traits', as they called them. They considered all of these to comprise varieties of human ability or capability, or, what I shall often call them here, 'skills'.

Later, E. L. Thorndike added to the classification of forms of intelligence when he proposed a special kind of "*social intelligence*"; he considered it distinct from traditionally measured intellectual intelligence, and identified it as "the ability to understand and manage people" (Thorndike, 1920, p. 275). Some decades later, social psychologists Bruner and Tagiuri (1954) directed attention to three distinguishable abilities involved in the perception of others: the ability to recognize the emotions that others feel based on their expressions, the ability to judge personality traits on the basis of the external behavior of others, and the ability to form impressions and formulate judgments of others. During the next decade, J. P. Guilford proposed a "multidimensional conception of intelligence" in terms of which he classified some 120 distinguishable abilities (Guilford, 1967, p. 467).

Later, Salovey and Mayer (1990), Gardner (1993/1983, 1989), and Goleman (1995) proposed the recognition of "*emotional intelligence*" defined in terms of a set of skills that enables one to become aware of one's own emotions and those of others. More recently, in the author's studies of the psychology of animal rights (Bartlett, 2002) and of human evil (Bartlett, 2005), I proposed the need to recognize "*moral intelligence*," defined in terms of a set of four specific, basic abilities that enable an individual to avoid succumbing to psychologically normal predispositions to violence, aggression, and destructiveness.

These historical examples drawn from slightly more than a century's research concerning human intelligence provide a useful background framework. We can see, as psychologists have become increasingly aware of distinguishable sets of human abilities, that they have often associated those

sets of abilities with corresponding varieties of intelligence. There are positive values in doing this, for it can help us to understand more clearly and explicitly different classes of human abilities; it makes it possible to focus investigation on identifiable discrete abilities; it may help us to learn how to develop and cultivate humanly important varieties of skills; and it enables us to understand how individual people vary in their capacities to learn and improve those skills.

In keeping with this gradually solidifying paradigm that links distinct sets of abilities with distinguishable forms of intelligence, toward the end of Bartlett (2005, p. 304), I coined the term ‘*epistemological intelligence*’ to refer to a set of special abilities, abilities that are involved in the study and development of epistemology when it is understood as a skill-based discipline. Like all abilities that are correlated with separately recognizable forms of intelligence, those that underlie epistemological intelligence are—as I have already suggested—highly variable among individuals. In fact, as we shall see, there are convincing reasons why the skills involved in epistemological intelligence are subject to so much individual variability.

#### **§4. Is epistemological intelligence no more than a theoretical construct?**

Definitions come in several kinds; the most familiar varieties include lexical definitions (as found in a dictionary), stipulative definitions, and real definitions.<sup>391</sup> When we question whether a definition—such as the definition of “epistemological intelligence” presented here—is a mere theoretical construct, at the root of the question is a concern whether what is being defined may be nothing more than a semi-arbitrary assemblage of defining factors (the “skills” I have associated with the term “epistemological intelligence”), which are gathered together and stipulatively asserted to constitute a distinguishable set that is noteworthy in some respect.

If we pause for a moment to consider this question and the concern it expresses, we find that, at its root, is a presumption that some definitions are, as it were, “privileged” in that they are thought to provide truthful information about the real world. Definitions of this kind have been given the name ‘real definitions’; they were described long ago by Aristotle as statements that express the essence of a thing—important facts about that thing, facts from which other truths can be obtained. Einstein’s definition of ‘simultaneity’ is

---

<sup>391</sup> For a detailed discussion of varieties of definition and an analysis of their principal uses and misuses, see Bartlett (2011, Chapter 2).



often thought to be of this kind. Unlike stipulative definitions, real definitions are thought to inform one about reality, and not simply to express a meaning that is to be agreed upon on the basis of mere convention. Real definitions purport to provide us with trustworthy information about real things by being descriptive of empirical reality, and therefore those definitions that do this successfully are considered to comprise true statements.

Elsewhere I have examined a closely related subject, the so-called ‘species problem’ in biology.<sup>392</sup> This is a famous problem for biologists and specifically for taxonomists. For centuries, there has been controversy over whether taxonomic divisions of classification are “discovered” or “imposed” upon the multiplicity of living things, whether they refer to “real divisions” or are fundamentally arbitrary.

We come face-to-face with the same issue in connection with any definition of a form of intelligence: Does that definition, one might ask, really discriminate reality along naturally existing lines of demarcation?, or Does that definition merely stipulate, perhaps on someone’s self-asserted authority, that it should be accepted as a convention, and possess no more compelling force than this? Let us call this the ‘definition problem’ in parallel with the ‘species problem’ of biology.

Without unreasonably broadening my focus here, I do not propose to develop a full answer to these questions, and, besides, interested readers can find that answer given elsewhere.<sup>393</sup> For purposes here, a short summary of the conclusions previously reached may be sufficient, translating those conclusions as they apply to our present discussion:

Those conclusions are, essentially, two: First, whatever efforts we make to answer the question whether “epistemological intelligence” constitutes a real definition of a “distinctly real variety of human intelligence”—or, instead, “merely a stipulated theoretical construct”—those efforts *must* fail because such a question does not recognize the theoretical impossibility of what it seeks: There is an inescapable ambiguity that necessarily affects *any* identification that clusters together a set of defining elements, as I will explain in a moment. And, second, any such clustering of defining elements presupposes a framework in terms of which those elements may be identified, and, by virtue of this inescapable fact, that *particular identified clustering of elements cannot not* be accepted without inconsistency: The very identification of the elements comprising the cluster is, in my terminology, “self-validating.”

To explain these claims as briefly as possible here:<sup>394</sup> When we consider

---

<sup>392</sup> Bartlett (2015).

<sup>393</sup> See the previous note.

<sup>394</sup> Further explanation and justification for these claims is given in Bartlett (2015).

any set of elements and recognize that a subset of them shares a certain property, the very possibility of this recognition relies upon the application, sometimes explicitly and sometimes not, of a criterion of commonality. It is only relative to such a criterion that a subset of elements can be recognized as sharing the property or properties specified by the criterion. The “inescapable ambiguity” mentioned in the previous paragraph results from the logical fact that even elements that we perceive to be different share the same number of properties as do elements that we perceive to share the same properties. Theoretical physicist and mathematician Satoru Watanabe was one of the first people to recognize and to prove this “logical fact” which he called ‘the theorem of the ugly duckling’.<sup>395</sup> In his words, the theorem “*claims that an ugly duckling and a swan are just as ‘similar’ to each other as are two swans*” (Watanabe, 1965, p. 39). Here lies the justification for the first claim relating to the inescapable ambiguity involved in any identified clustering of elements that we may recognize.

Once a criterion of commonality has been established, relative to which we recognize that, from a set of elements, a certain subset or cluster satisfies that criterion, there is no possible contention that *can* arise without inconsistency. —Notice that the preceding sentence begins with the establishment of a “ground rule”: *Given* a certain criterion of commonality, and *given* that we do recognize that a certain group of elements satisfies that criterion, on this established basis *then* to reject the recognized cluster—which we have identified through the application of the specified criterion—would be inconsistent. It is inconsistent in a particular way, one that attempts to “pull the carpet out from beneath our feet”—to deny, in other words, the very preconditions that must be met in order to identify the specified cluster. The projective variety of self-undermining inconsistency that is involved is now familiar to readers of this book.

The conclusion that follows is that we would be engaged in a theoretically futile search if we insist upon deciding between the two alternatives, whether epistemological intelligence is a real definition or a theoretical construct. Epistemological intelligence, as defined here, is a recognizable clustering of certain skills, and once identified as it has been defined, its value as an identifiable set of abilities is wholly a matter of choice. This applies equally to all defined varieties of intelligence, whether the skills they cluster together are evidenced through IQ tests, memory tests, dexterity tests, moral reasoning tests, emotional or aesthetic sensitivity tests, etc. Each and every identified variety of intelligence is a matter of choice, which is a function of the impor-

---

<sup>395</sup> A detailed account of his theorem is given in Bartlett (2015). See especially Watanabe (1965); then Watanabe (1969, 1985, 1986).

tance placed on the relevant underlying skills. In the case of epistemological intelligence, that choice depends upon the usefulness to us in recognizing that a certain cluster of skills is important and fundamental in epistemological thinking. The concept of epistemological intelligence can be useful if one's interest in epistemology includes the development and improvement of the set of eight skills previously described; otherwise, it probably is not.

### §5. The psychology of philosophers

In the first section of this appendix, I raised two questions: Both suggest that when the eight skills I have associated with epistemological intelligence are made an area of study and cultivation, a large amount of individual variability makes itself evident: Some students of philosophy, as well as some professional philosophers, exhibit these skills in a strongly evident way, while others do not. The two questions I raised were: Why is this the case? and What does this reveal about the workings of the individual philosophizing mind? If one is interested in psychological dimensions of philosophizing and in the psychology of professional philosophers, these two questions can be illuminating. An interest in the nature of epistemological intelligence and in ways in which such intelligence can be developed and cultivated, leads us to confront these questions.

With training and research in clinical psychology, I became interested in this general area of study, the psychology of philosophy and of philosophers, and published a group of papers in the 1980s dealing with this topic.<sup>396</sup> Since then, as far as I've been able to determine, philosophers have not added to this research about their own psychology, very likely because few have training in clinical psychology. In the meantime, however, psychology and its recent offshoot, behavioral economics, have made relevant observations about a similar underlying psychology possessed by any group of people who defend vested interests. Given philosophy's fundamental credo that an unexamined life—even one that remains unexamined from a psychological perspective—is undesirable, in what follows I shall try to summarize some of these main psychological results that apply specifically to the philosophical enterprise and to the individual psychology of many philosophers and of many students attracted to philosophical study. By doing this, we shall be able to throw some light on the specific skills I have clustered under the heading of epistemological intelligence, and explain the wide variability of those skills among individuals.

---

<sup>396</sup> Bartlett (1986a, 1986b, 1989).

I start with an abbreviated summary of the principal psychological observations made in the three published papers cited in the last footnote. Some of this information overlaps with and adds to earlier discussions in {1} and {2}.

In Bartlett (1986a), I directed attention to several forms of defensiveness exhibited by many philosophers:

- i. Controversy between philosophers ensues when their systems of belief come into conflict. For many, their preferred systems of belief are rooted in their sense of self and of professional identity, and therefore philosophical disagreement can easily be experienced as personally threatening. Challenges to that set of beliefs can be intimidating and can often result in a defensive response.
- ii. When a system of beliefs defines one's sense of identity, that system of beliefs tends to close the mind to competing, alternative, or conflicting ideas. In this sense, a philosophical system of beliefs becomes an ideology, allegiance to which comes to possess the character of a solipsistic creed, hermetically closing off that system of beliefs from potential challenge.
- iii. The close wedding of personal and professional identity to a self-enclosed, hermetically sealed system of ideas results in an inevitable filtering effect, one that sieves incoming, as well as outgoing, would-be communications with other minds. This filter imposes restrictive limits on what one is willing to see, admit, or question.
- iv. Philosophical positions and ideologies generally—like their proponents—are exceptionally resistant to external criticism; they possess an “almost inconceivable hardihood” (Passmore, 1961, p. 63), a conceptual and psychological imperviousness in the face of opposition.
- v. A philosopher's sense of self, reflected in his or her system of intellectual and psychological commitments, is frequently expressed in an intention to formulate an all-inclusive theory of reality, and hence to claim that it “includes all relevant evidence.” This is a frame of reference that is, essentially, exclusionary.
- vi. As a result, philosophical positions and the philosophers who espouse them tend to become monadically isolated from one another, preventing genuine communication.

- vii. As long as (i)–(vi) remain dominant in the psychology of philosophers, explicit and genuine controversy between competing views, true dialogue between conflicting ideologies, cannot succeed.

In Bartlett (1986b), I summarized the principal set of defining characteristics of clinical narcissism as it relates to the psychology of philosophers. The following caveat was expressed early in that paper:

Philosophers, like other people, are subject to human frailties. Some are probably clinical narcissists. I do not know if a larger proportion of philosophers is narcissistic than are theologians, poets, composers, artists, or writers. But probably, for reasons I will try to make clear, a greater proportion of the philosophical population suffers from characteristics of unacknowledged narcissism than do, for example, scientists... [T]he nature of philosophic activity promotes and is encouraged by many qualities of personality which closely resemble qualities that define narcissism. (Bartlett, 1986b, pp. 21-22)

The value in recognizing aspects of clinical narcissism in the psychological constitution of many philosophers is not to derogate, but rather to understand the challenges that result in a profession which—like all professions, as the interest inventories of psychometrists show—possesses among many of its practitioners a distinctive, recognizable, and characteristic set of psychological propensities. As described in Bartlett (1986b), certain of the defining characteristics of clinical narcissism apply to the philosophical personality:

- viii. During its long history, philosophy and its practitioners have generally resisted the development of a unitary, evaluative framework. In contrast, science is the enemy of intellectual narcissism: It does not sanction the privilege of privately formed judgment, but requires the consensus, based on evidence, of the scientific community. Occasionally, individual philosophers have urged that philosophy become scientific, which is equivalent to advocating that the private willfulness of individual systems-building submit itself to an agreed-upon unitary methodology.<sup>397</sup>

---

<sup>397</sup> As Husserl once commented, “philosophers meet but, unfortunately, not the philosophies. The philosophies lack the unity of a mental space in which they might exist for and act on one another” (Husserl, 1965, p. 5).

As we have seen in (i)–(vi) above, there are compelling psychologically based forces that stand in the way of such a development.

- ix. The psychology of philosophers is characterized by contentiousness: It is the life-blood of position-taking and resulting philosophical interaction. Its goal is to show that one is right and that the other is wrong. As a result, philosophical argumentation combines self-demonstration with attempts to show that the opposition is mistaken.
- x. Philosophical position-taking is essentially defensive, seeking to evade criticism. Often this is attempted through the cultivation of vagueness and terminological obscurity, a “smoke-screen in which only initiates can navigate confidently” (Bartlett, 1986b, p. 25). I have called this propensity ‘intellectual deviousness’.
- xi. The psychological underpinning of the practice of a discipline that defines itself through position-taking takes the form of resistance to external pressures that would force change in propounded positions. This internal resistance of a discipline, itself a kind of “intellectual lassitude” that develops within the mind of the position-taking philosopher and is then incarnated in the position espoused, is one of the defining features of clinical narcissism. It serves to mask fear that one’s efforts are tenuous or flawed.

In Bartlett (1989), I sought to bring together certain of the preceding psychological observations within a context that recognizes that all disciplines—some to a greater and some to a lesser extent—attract and encourage practitioners who have distinguishable personality structures. In connection with mathematics, for example, an obsessive-compulsive personality structure can sometimes confer a significant benefit upon creative work. Outer-oriented, socially directed individuals are frequently attracted to sales, teaching, or social work, and can become competent professionals in these fields. There is sometimes a positive role which a professional’s personality structure contributes to his or her growth, and also contributes to the development of the person’s chosen profession. But sometimes the characteristic personality structure possessed by practitioners of a discipline can interfere with and obstruct its positive growth.

The characteristics and propensities listed above, (i–xi), may, whether all or only a subset of them, be combined within a single philosophical

personality. When this happens, the person's sense of self and professional identity can lead him or her to embrace positions that come to function as ideologies; the filtering effect that results encourages conceptual and psychological imperviousness to outside criticism, an imperviousness that is fundamentally exclusionary, monadically isolated, and resistant to the development of a broader, embracing, unitary evaluative framework. The mindset to which these propensities lead is one of contentiousness, defensiveness, and resistance to pressures that would force a change of view.

After nearly three decades of still continuing psychologically focused observations of my fellow philosophers and of students who are drawn to philosophy, the conclusion I reached in the late 1980s remains one that I have found no evidence to revise:

When these qualities [i.e., (i-xi)] dominate a field of thought, they are surely maladaptive. The discipline cannot advance. It chases its tail. The clutch slips, and though the engine races, the vehicle that philosophical reflection affords remains stationary. It will continue to make promises, but remain incapacitated, unable to fulfill them.... If the psychological observations offered here are on the right track, then it follows that it will be difficult, even *a priori* impossible, to find anything that is capable of conveying an effective message to the majority of philosophers across the boundaries of their island universes of meaning. To be enmeshed in an ideology of one's own fashioning is incompatible with the awareness that one's commitments are ideological. (Bartlett, 1989, pp. 304-305)

### **§6. The personality structure of philosophers as seen through the lens of confirmation bias**

Confirmation bias, as the term is typically used in the psychological literature, connotes the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand.... If one were to attempt to identify a single problematic aspect of human reasoning that deserves attention above all others, the confirmation bias would have to be among the candidates for consideration. Many have written about this bias, and it appears to be sufficiently strong and pervasive that one is led to wonder whether

the bias, by itself, might account for a significant fraction of the disputes, altercations, and misunderstandings that occur among individuals, groups, and nations.

– Raymond S. Nickerson (1998, p. 175)

In Bartlett (1989), I commented: “To my knowledge, no experimental psychological study has ever been made of philosophers as a group” (p. 297). In the years since then, as far as I’ve been able to determine, it is still the case that no such experimental psychological study has been undertaken. Nonetheless, two areas of research have developed whose findings are indirectly applicable to the focus here on what I take to be the dominant personality structure of philosophers. I shall look at certain of the relevant results of one of these areas of study in this section, and consider the second area of research in the following section.

In the past few decades, psychologists and behavioral economists have become interested in investigating the human phenomenon of *confirmation bias*. So much has now been written about it that confirmation bias has received the pet name ‘myside bias’. There are many studies, both experimental and theoretical, that could be mentioned here and applied to the personality structure of philosophers and of students who gravitate to the profession. However, my assessment is that only rather little that has been written and experimentally evidenced having to do with confirmation bias adds appreciably to the philosophically focused observations already described in §5. But sometimes a slight change of perspective can bring with it a stronger understanding of a subject-matter; to that end, I summarize a few of the conclusions that psychologists and behavioral economists have come to affirm that are pertinent to the previously described psychological propensities of philosophical defensiveness (i–vi) and of characteristics of philosophical narcissism (viii–xi). By doing this, we may gain a more complete understanding of some of the major blocks that stand in the way of many individuals who might wish to develop and improve the specific epistemological skills (1–8) identified in §2.

One of the most widely accepted conclusions about confirmation bias that has now earned near-universal consensus among myside researchers is this: Once a person has taken a position, the immediate psychological sequelae are to find ways to justify it. To give an example that closely parallels the position-taking of philosophers, consider taxonomists.<sup>398</sup> A taxonomist’s focal interest is to formulate a system of classification that purports—whether in

---

<sup>398</sup> See, e.g., Nickerson (1998, pp. 183–184).



biology, botany, atomic physics, materials science, etc.—to express *the real structure*, the authentic, empirically based divisions and distinctions, of the domain of objects of concern to the taxonomist. Having constructed such a taxonomy, the natural, and hard-to-resist, psychological consequence for the taxonomist is the strong tendency then to perceive his or her subject-matter *in terms of* the specified system of classification. This effect upon the conceptual-perception faculties of the taxonomist is an almost direct embodiment of the Sapir-Whorf linguistic relativity hypothesis, for readers acquainted with that thesis.

And so, the first consequence of the creation of a taxonomic system of classification is to *channel* the taxonomist's own thinking and observations so as to conform to the designated categories. But, closely following on this is a second important consequence, that of leading the taxonomist actively both to search for substantiating evidence that supports the now-preferred system of classification, and, unfortunately, also to avoid considering evidence that conflicts with or undermines that system. Both selective, cherry-picking-searching and the exclusion of contrary evidence are instances of confirmation bias.

The two psychological sequelae in this taxonomy-creation example—channeled, filtered thinking and biased evidence selection—which have repeatedly been observed in studies of confirmation bias, are recognized to involve both a motivational component as well as a component reflecting cognitive limitations,<sup>399</sup> some of which are discipline-specific, and some quite general. It is admittedly often difficult to disentangle which component is a stronger contributor to confirmation bias; some researchers lean more one way, some the other. But in the context of our present concern with the dominant personality structure of philosophers, this difficulty is significantly reduced because it is reasonably straight-forward to separate the motivational component from the cognitive skills component when we consider the two previous lists: the list of the eight identified epistemological skills and their associated motivational factors, and the list of the eleven characteristics of philosophical defensiveness and narcissism. We shall return to this subject when we examine the reasons for individual differences in epistemological intelligence.

### §7. Non-philosophical studies of so-called 'epistemological understanding'

A second area of research has developed in the past several decades; its contributors have largely been professors of education and professors of the psy-

---

<sup>399</sup> See, for example, Kunda (1990) and numerous publications that have followed her paper.

chology of education. Their work has some indirect applicability to an analysis of the skills that constitute epistemological intelligence and deserves brief mention here. These researchers have attempted to relate what they have called ‘epistemological understanding’ to elementary and high school student skills in argumentation.<sup>400</sup> The notion they employ of “epistemological understanding” is rather general and is loosely defined to include four basic aspects of knowing: degree of certainty, simplicity, justification, and source.<sup>401</sup>

Admittedly, their use of the term ‘epistemology’ has become very casual and semantically undisciplined, and would satisfy few philosophers whose expertise lies in the discipline of epistemology. From the standpoint of technically oriented philosophical epistemologists, the “epistemological understanding” that has interested these authors is situated on a comparatively elementary level, both in terms of the skills involved and the levels of student education that are investigated. Even so, this non-philosophical area of study is one of the few that partially overlaps the interest here in examining the set of skills associated with epistemological intelligence.

The work of these psychologically focused academicians has tended to endorse a developmental model of cognitive growth, originally proposed by Kuhn (1999, 2000). She suggested that “epistemological understanding” proceeds in a series of stages from childhood to adolescence to adulthood: Childhood absolutism is characterized by the belief that knowledge is definitive, unambiguous, directly based on experience of reality and supported by the imprimatur of authorities. In adolescence, this stage of absolutism gives way to a “multiplist” view that sees knowledge as ambiguous, subjectivist, and relative. And then, by the time adolescents reach early adulthood, they ideally begin to develop a capacity to evaluate knowledge claims critically and to accept the existence of unitary norms of justification and inquiry.

Such an idealized stages-of-development model of the gradual growth of “epistemological understanding” is worth mentioning in our present discussion if only because it parallels on a rudimentary level (of education and of cognitive skills) what we have explicitly in view here. However, although it can be of interest in its own right, a fundamentally questionable aspect of this research is the widely shared belief among many of its researchers that “epistemological understanding” should be considered in “dispositional” terms rather than in terms of an individual’s actual competence or general intelligence—meaning that the wide divergence in student performance in tasks that require them to justify basic knowledge claims calls upon their “dispositions”

---

<sup>400</sup> Cf., e.g., Weinstock, Neuman, & Tabak (2004), Mason & Scirica (2006), Weinstock (2009).

<sup>401</sup> Suggested by Hofer & Pintrich (1970) and often, with some variations, followed since then, as, for example, by Weinstock (2009).

to think in a flexible way and to be willing and able to change their beliefs once they encounter contradictory evidence.<sup>402</sup>

To refer to these mental proficiencies (my attempt to use a neutral term) by using the word ‘dispositions’ is, I think, to suggest that individuals can, at least sometimes, voluntarily change their cognitive behavior provided they are so *disposed*. I suspect that here the use of the word ‘disposition’ may bow to political correctness in order to direct attention away from, or even to deny, *individual cognitive differences*. To my knowledge, however, it has not been shown that the mental proficiencies in question are, as the term ‘dispositional’ seems to suggest, of a kind that the individual is able, in any sort of comparatively free or unrestricted manner, to change his or her cognitive behavior. These mental proficiencies are, in fact, mental skills that are not equally apportioned or endowed, as I shall try to make clear in what follows. Whether they can be taught or developed through individual initiative may, as we shall see, be open to some doubt.

...

To summarize this and the previous section:

The findings of studies of confirmation bias that relate most directly to the philosophical personality are these: Position-taking—of any kind, whether philosophical or not—has a psychological dynamic for which there is now strong evidence. It constitutes a dynamic—i.e., it forms an interconnected system—which involves the interplay of the following psychological forces: Once a person takes a position, his or her immediate, virtually automatic psychological tendency is (a) to seek ways to justify it, (b) to perceive reality in conformity with that position, (c) to filter his or her thinking in the terms of the position taken, and (d) to engage in biased evidence-selection, choosing evidence that favors the position, and excluding evidence that does not.

The second area of research that has focused on so-called ‘epistemological understanding’ in students has suggested that as people age, they undergo progressive stages of development in which they, ideally, become increasingly capable of critical thinking in their evaluation of claims to knowledge. To what extent this capacity can be developed through training, to what extent naturally occurring individual differences contribute or interfere, and whether or not individual differences in these proficiencies are “dispositions” or instead comprise “skills” (in the sense of competencies or expressions of in-

---

<sup>402</sup> Stanovich (1999), Stanovich & West (1997, 1998), Weinstock (2003).

telligence)—all of these pose questions whose answers have not been convincingly established.

Despite important unanswered questions, there are obvious direct applications to the psychology of philosophers of the above findings (a–d) and of the suggestion that individual proficiencies in evaluating knowledge claims may be developmental, and perhaps also subject to individual differences.

To complete this discussion of confirmation bias in its relation to the psychological profile of philosophers, it will be useful to refer to the list of psychological characteristics given earlier in §5. For the reader’s convenience in referring to that list, it is reproduced below in somewhat abbreviated form:

- |       |   |
|-------|---|
| (i)   | Preferred philosophical systems of belief are often rooted in a philosopher’s sense of self and professional identity; philosophical disagreement can easily be experienced as personally threatening.                                  |
| (ii)  | When a system of belief is linked to one’s sense of identity, it can close the mind to competing, alternative, or conflicting ideas, and can become a solipsistic ideology, closing off that system of belief from potential challenge. |
| (iii) | The result of (i) and (ii) has a filtering effect, which sieves incoming, as well as outgoing, would-be communications with others, imposing limits on what one is willing to see, admit, or question.                                  |
| (iv)  | Philosophical positions, and ideologies generally—like their propounders—become exceptionally resistant to external criticism.  |
| (v)   | A philosopher’s system of intellectual and psychological commitments often presumes to be an all-inclusive theory of reality, to “include all relevant evidence,” thereby becoming essentially exclusionary.                            |
| (vi)  | As a result, philosophical positions, and the philosophers who espouse them, tend to become monadically isolated from one another, blocking genuine communication.  |
| (vii) | As long as (i)–(vi) dominate the psychology of philosophers, genuine controversy and dialogue between competing views cannot succeed.   |

- (viii) As a consequence of the foregoing, philosophy, unlike science, has resisted the development of a unitary framework and methodology. Philosophers who have urged the adoption of a shared standpoint and method implicitly advocate abandoning the private willfulness of individual systems-building. But (i)–(vi) are compelling psychological forces that stand in the way of doing this.
- (ix) The psychology of philosophers is characterized by contentiousness, the goal of which is to show that one is right and that the other is wrong.
- (x) Philosophical position-taking is essentially defensive, seeking to evade criticism, often by means of vagueness, terminological obscurity, and self-insulation from outsiders.
- (xi) A basic psychological underpinning of philosophy is resistance to external pressures that would force change in positions taken; this resistance comprises a kind of “intellectual lassitude” that can mask fear that one’s efforts are tenuous or flawed.

### **Appendix II. Table 1. Common psychological traits of philosophers**

The main findings relating to confirmation bias that have direct application to the philosophical personality are, then, these: Once a person takes a position, his or her immediate, virtually automatic psychological tendency is

- (a) to seek ways to justify it,
- (b) to perceive reality in conformity with that position,
- (c) to filter his or her thinking by virtue of the position taken, and
- (d) to engage in biased evidence-selection, choosing evidence that favors the position, and excluding what does not.

There are clear-cut correlations between these four propensities and the above combined list of psychological characteristics listed in Table 1:

The drive to find ways to justify a position that one has taken, (a) above, is strongly expressed in the contentious psychological propensity of philosophers to show that one is right and that the other is wrong (ix).

The next two psychologically compelling effects of confirmation bias—to perceive reality, (b), and to filter one’s thinking, (c), in conformity with a position that one has taken—take the form in the philosophical personality of filtering incoming, as well as outgoing, would-be communications with others; the filter that is established imposes restrictive limits on what one is willing to see, admit, or question (iii).

The last, and most widely recognized effect of confirmation bias, the selection and rejection of evidence to suit the position taken, (d) above, takes the form in the philosophical personality of a self-enclosed, self-isolating, exclusionary system of intellectual and psychological commitments that claim that the philosophical position endorsed includes all relevant evidence (v).

The philosophical characteristics (iii), (v), and (ix) embody, then, the principal relevant results of studies of confirmation bias. The remaining characteristics (i, ii, iv, vi, viii, x, and xi) complete a description of the psychological profile of philosophers which, while sharing several features of the general psychology of confirmation bias found in all position-taking, incorporates a group of further important characteristics.

The single “meta-level” property, that of combining characteristics (i–vi) within a single personality structure (vii), is significant, since it points to the claim implicit in this appendix, that it is by virtue of the combination of the various propensities (i–vi and viii–xi) that a total, integrated psychological profile comes to be. The varied psychological characteristics I have listed function in *synergy*: They interact and mutually reinforce one another so as to produce a combined total effect that is psychologically distinctive and of central relevance if philosophy is to achieve its goal of clear and dispassionate self-examination.

The above psychologically based observations are of potential value to practitioners of philosophy for two principal reasons: They can provide a framework for honest psychological self-examination and growth, and not a depressing dead-end; and they help us to understand important psychological factors that must enter into our understanding of epistemological intelligence, to which subject I now return.

## §8. Epistemological intelligence and individual differences

I know that most men—not only those considered clever, but even those who are very clever, and capable of understanding most difficult scientific, mathematical, or philosophic problems—can very seldom discern even the simplest and most obvious truth if it be such as to oblige them to admit the fal-

sity of conclusions they have formed, perhaps with much difficulty—conclusions of which they are proud, which they have taught to others, and on which they have built their lives.

– Leo Tolstoy (1899, p. 124)

It should by now, I hope, strike most readers that Table 1's set of eleven psychological propensities, when these operate in synergy within the individual philosophizing mind, are essentially undesirable because they are *limitative*. That is, they stand in the way of an individual's intellectual as well as personal growth, and when they represent the dominant psychological profile of a profession, they similarly stand in the way of its development.

As I have emphasized in previous publications relating to the psychology of philosophers,<sup>403</sup> observations such as these are not intended, nor should they be taken, as disparaging of the discipline, as derogatory, negativist, or an expression of hopelessness. I hold certain of the core objectives of philosophy in high esteem and have sought to contribute to the discipline in ways that enable it to overcome its undesirable *internal limitations*. The observations I have attempted to communicate seeks to serve as stepping stones to a psychologically comprehensive understanding of the phenomenon of philosophy, an understanding that may help to advance its development.

In §2, I identified a set of eight basic skills that I associate with epistemological intelligence (again, for the reader's convenience, that list is reproduced here):

- (1) a commitment to logical coherence and alertness to inconsistency
- (2) a commitment and a sensitivity to identify beliefs that are baseless, that is, which have no empirical or logical justification
- (3) a commitment and a will to eliminate, to the extent that this is individually possible, baseless beliefs from one's own thinking
- (4) a commitment to seek revisionary replacements, to the extent that this is conceptually possible, for beliefs recognized as baseless, and a determination to accomplish this

---

<sup>403</sup> See note 396.

- (5) an ability to recognize that all claims to knowledge are inextricably framework-relative, that is, cannot be asserted when disassociated from the conceptual framework(s) that must be accepted in order for it to be possible to make those claims
- (6) a heightened awareness of the ways in which some scientific and many commonly accepted and widely used concepts trespass beyond the frameworks they presuppose, and by doing this, become self-destructively incoherent, i.e., metalogically projective
- (7) a “mental dynamic” of a certain sort, one that has been developed within the individual’s mind, perhaps through training or perhaps simply because the person’s mind finds it natural and desirable to invest credence only in rationally justifiable ways of understanding and rationally justifiable claims that can be made on this basis
- (8) the establishment of what I have called a ‘*rational bridge*’ which connects the preceding reflectively, self-consciously justified and formulated understanding and claims with both the individual’s own predispositions to behave in certain ways, and actual conduct that is consistent with those predispositions

**Appendix II. Table 2. Basic skills associated  
with epistemological intelligence**

My purpose in this section is to look more closely at the eight basic epistemological skills shown above in Table 2 by considering their relationships both with the specific forms of defensiveness exhibited by many philosophers and with the defining characteristics of clinical narcissism that apply to the philosophical personality which were listed in Table 1. As we shall see, the degree of the strength of such relationships can facilitate or obstruct the attainment of epistemological skills when these are sought within a psychological framework characterized by these kinds of defensiveness and expressions of narcissism. These relationships will explain in large measure the variability that we see in the extent to which individuals can acquire and cultivate those skills.

In a different context, I described the nature of a “closed system of thought” in the following terms:

A closed system of thought capitalizes on the vicious internal circularity of its method of responding to outside challenges.



Thanks to its circular logic, “[t]he true believer ... can prove to his satisfaction everything he believes, and he believes everything he can prove” (Koestler, 1967, p. 290). For the believer, his understanding of the world, his emotional approach to it, and the behavior he is motivated to engage in, often to the point of death, all of these are built upon a pathological way of interpreting experience. Experience is, one might say, remapped by the following features: (a) the closed system’s immunity to conflicting evidence, that is, its self-insulating ability to seal itself off from opposition; (b) the unbridgeable separation of human reason and emotion; (c) the closed system’s universalizability; [and] (d) irrefutability.... Together, these forces support and maintain a cognitive dynamic that reinforces the species’ emotion-dominated response to any form of opposition that questions the needs and interests of the closed system. And many human cognitive frameworks are closed in this way: an ideology, a nation’s or a corporation’s self-interested policies, a body of religious dogma, any group’s self-centered manifesto, or generally, homocentric selfishness. (Bartlett, 2008, p. 358).

It should be evident that the psychological limitations listed in Table 1 describe a manifestly closed system of thought, one which, if my observations are correct, characterizes the dominant psychology of philosophers and many of its students. It should also be evident that this closed system of philosophical thought is, on a fundamental psychological level, *incompatible* with the essentially open system of thought that is established when all of the skills enumerated in Table 2 are fully functional in an individual mind. Not only are these two sets of propensities incompatible, they are *opposed* to one another. A mind that possesses the skills listed in Table 2 is “immunized,” so to speak, from the self-limiting propensities identified in Table 1. To express this concisely: *Epistemological intelligence inhibits the formation of characteristic philosophical defenses as well as philosophical expressions of ideological and intellectual narcissism.*

The topic of individual differences has unfortunately, in many applications, become too politically hot for many researchers to handle. We certainly are willing to recognize that at least some human abilities vary widely—that the skills of genius, for example, are not conferred on all equally. Despite the obviousness of this fact, there is a pronounced wish among many researchers today to avoid any reference to cognitive skills that are unequally apportioned

in the population. Some of this evasion may be explained because, to accept individual differences, especially individual cognitive differences, is to accept some minimal form of “discrimination,” even if this minimal form amounts to no more than the ability to *discriminate*—that is, to perceive and to recognize—more highly developed skills as differentially distinct from those less developed. The rest of the evasion is explained by a mistaken equation of democratic equality of rights with equality of abilities. From the perspective of any teacher wishing to encourage in his or her students the development of any set of skills, discrimination in the minimal sense mentioned above is essential, for that discrimination is an expression of the sensibility needed to recognize an ideal, an ideal in relation to which the educator’s teaching objectives are determined.

This is equally true when the educational goal is to help students develop the skills associated with epistemological intelligence. The skills of epistemological intelligence as they have been defined here are capable of attainment to different degrees by different minds. It will not be hard to see why this is the case.

Before proceeding, I want to introduce the reader to what I have called the ‘Romanes Principle’.<sup>404</sup> The psychologically focused observations presented here, made of my philosophical colleagues and students of philosophy over a period of decades, are phenomenologically and clinically descriptive and are, at this point, not yet supported by experimental or psychometric studies. This fact does not sit well in the context of the prevailing strong distrust of individual diagnostic judgment when not backed by double-blinded, statistically sound, and replicated experimental studies. In relation to that distrust of individual judgment, the Romanes Principle expresses a viewpoint which today cannot help but appear heretical. However, although there is an important role in science for caution, there is no place in it for thick-headedness. Nineteenth century psychologist George Romanes did not devalue experimental evidence, but he was level-headed:

[N]o one can have a more lively appreciation than myself of the supreme importance of experimental or historical verification, in all cases where the possibility of such verification is attainable. But in cases where such verification is not attainable, what are we to do? We may either neglect to investigate the subject at all, or we may do our best to investigate it by employing the only means of investigation which are at our disposal.... [I]n the science of psychology, nearly all the con-

---

<sup>404</sup> Bartlett (2011, pp. 203-204). See also Bartlett (2005, pp. 226-227)

siderable advances which have been made, have been made, not by experiment, but by observing mental phenomena and reasoning from these phenomena deductively. (Romanes, 1895/1883, p.12)

Once reasoning based on observation does provide a demonstration of a certain fact, we will sometimes find that “the proof itself is too complete to admit of any question.” (Romanes, 1889, p. 400)

There are, Romanes suggested, facts we can come to know about the world, and specifically about psychology, that do not rest on or require experimental verification. In the case of philosophy’s psychological underpinnings, I will go out on a non-experimental limb, affirm the Romanes Principle, and claim that we do not need to cringle in a state of dependency upon as yet unperformed psychological experiments that would verify what the history of philosophy already abundantly confirms. There is a place for phenomenological description that does not require experimental, statistical buttressing.

Once a basis has been established in phenomenological description, there is also a place for plausible reasoning. I now call on the reader to accompany me in taking a few plausible steps:

There are psychologically simple reasons why many people find the skills associated with epistemological intelligence to be challenging and often emotionally and intellectually *inapplicable* to their outlooks and *unacceptable* to their persons. Table 1 lists a group of these reasons, reasons that especially affect philosophers. We are, I submit, confronted by *psychological facts* which are borne out by philosophy’s long history and by the psychological profile of the majority of philosophers and students who are attracted to philosophical study. But it is important to recognize that these are not psychological facts unique to philosophers and students of philosophy, but are facts strongly supported by evidence from clinical psychology, psychiatry, ethology, and quantitative history that relate to the *normal human psychological constitution* at this time in our species’ development.<sup>405</sup> In many respects relevant to our analysis here, philosophers and the average, philosophically untrained, psychologically normal population are similarly limited by the dynamics of their own systems of belief.

We may sum up what most plausibly happens when, to use a mechanical

---

<sup>405</sup> For a detailed study in support of this statement, see Bartlett (2011, esp. Chap. 9 and Appendix III, “The Distribution of Mental Health”), and Bartlett (2005).

analogy, a thinker's mental "gears"—consisting of the philosophical forms of defensiveness and narcissism listed in Table 1—are engaged in pursuit of the ideal goal defined by the skills fundamental to epistemological intelligence listed in Table 2: The gears will grind, will refuse to mesh, and will resist forward motion. A thinker whose mind is characterized by the forms of defensiveness and narcissism listed in Table 1 will experience deeply rooted recalcitrance when faced with the challenge, or with the need, to revise or relinquish beliefs that are emotionally or intellectually gratifying. The interplay of the forces of defensiveness and narcissism resists the formation and cultivation of epistemological skills.

Referring once again to Table 2, both the general public as well as many professionally trained philosophers and their students frequently experience great reluctance to give up baseless, unjustified beliefs (3); they are able to summon negligible determination to seek revisionary replacements for such baseless beliefs (4); and they balk at the injunction not to trespass beyond presupposed frameworks of reference (6), for that injunction is mightily resisted by a mind that is accustomed to, yearns for, and is compelled to seek horizon-transcending, projective access to supposed truths and corresponding realities that transcend the very possibility of access to them.

At the same time, comparatively few individuals possess a psychological constitution that limits their beliefs to rationally justifiable understanding and rationally justifiable claims (7), and, in my own observations, even fewer people possess minds in which a "rational bridge" exists to connect reflective, rational understanding and claims with actual cognitive and behavioral conduct consistent with their rationality (8).

When the skills I have associated with epistemological intelligence are sought, the mind encumbered by the limitations identified in Table 1 experiences the unavoidable consequences of the limitative psychological and cognitive shortcomings we have identified. To use a mechanical analogy again, the accelerator is pressed with the brakes fully engaged.

On the positive side, the skills associated with logical coherence (1), a capacity to become aware of baseless beliefs (2), and the ability to recognize the framework relativity of claims (5) are skills whose formation and development the majority tends to resist less: It is not uncommon for many students to be able to develop these abilities. But "resisting less" does not by any means entail that the corresponding *motivational* components will be present. The skill of being *alert* to inconsistency (1), the *commitment* to identify baseless beliefs (2), and the *determination* to exercise the ability to recognize framework relativity (5), these all involve active determination and resolve,

and here intellectual lassitude or lethargy (Table 1, xi) tends to play a dominant limitative role.

### **§9. Can epistemological intelligence be taught? Can it be learned?**

The first goal in studying any specific variety of intelligence is to understand the set of skills that make it possible. Once that understanding is attained, *how* to implement that understanding to make the skills associated with that form of intelligence more accessible and more achievable for people is a more distant goal. But the understanding must come first.

In §3, I referred to my study of the variety of intelligence that I have called ‘moral intelligence’. Related to this work, a number of researchers have examined what they believe are distinguishable developmental levels of “moral reasoning,”<sup>406</sup> much like levels of “epistemological understanding” are described by educators and psychologists of education. In the case of moral reasoning, there is unfortunately scant evidence (in fact I know of none) to show that the level of an individual’s moral reasoning is correspondingly reflected in his or her actual moral behavior in real life. To be “morally intelligent” requires the existence of a “moral bridge” that connects moral reasoning with actual behavior consistent with that reasoning; such a bridge is a comparative rarity in the psychologically normal population.<sup>407</sup>

A parallel issue concerns the unreliable to non-existent “rational bridge” in many people (Table 2, (8)). There is no question that many individuals can, through effective training, improve their reasoning skills. Not only this, but there is some evidence that some of the skills associated with general intelligence can be improved through teaching that specifically focuses on the development of problem-solving skills.<sup>408</sup> There is again, however, as in the case of moral reasoning skills, a paucity of evidence to show that general reasoning skills carry over to an individual’s subsequent rational thinking and behavior. Here, too, a “bridge” between developed intellectual skills and later cognitive

---

<sup>406</sup> For references to the literature, see Bartlett (2005, Chap. 18 and *passim*).

<sup>407</sup> Again, see Bartlett (2011, “The Distribution of Mental Health,” pp. 273-276).

<sup>408</sup> This area of investigation remains controversial. To cite evidence here only from the author’s own research: With grant support principally from the Lilly Endowment, at Saint Louis University during the years 1976-83, I designed and regularly offered a campus-wide course, *Patterns of Problem Solving*, that provided students with training in a wide range of problem-solving skills. Pre- and post-testing of students enrolled in this class showed significant improvements in IQ as measured by the California Test of Mental Maturity (see Bartlett (1976-77)). Students whose IQs were measured could not, however, be followed longitudinally in later years to determine whether their IQ gains were long-lasting or short-lived.

and behavioral conduct is, at best, shaky.

The two “bridges” I have described would ideally link, in the one case, moral reasoning to moral behavior, and in the other, reasoning skills to subsequent rational thought and behavior.

What is needed is an “affective cement” that welds moral perception and reasoning to behavior. Conviction, as commonly understood, serves this connective function; it acts as the bridge that unifies an individual’s judgment, reasoning, and beliefs with behavior that conforms to them. Without conviction, reasoning falls on deaf ears; without conviction, there is no need for moral or reflective judgment to be expressed through behavior in real life. Seen in this way, conviction is a genuine ability, an ability to connect strongly felt emotions with behavior consonant with them. It is what enables someone who is persuaded by reflective thought or by moral principle to act consistently with that thought or principle. (Bartlett, 2005, p. 280)

This manner of speaking is evidently impressionistic; it makes recourse to analogy and metaphor, and is imprecise clearly because of a lack of empirical data in this area. Notwithstanding the Romanes Principle, my commentary here is forced to be objectionably metaphorical. —Of course there are no such actual “bridges,” though perhaps neuroanatomists may in time discover neural structures or patterns of organization in the human brain that provide these figures of speech with a physiological basis.<sup>409</sup>

Although stated informally and figuratively, the point here, however, will, I hope, not be lost on the reader—that the epistemological skills I have identified require just the kind of “cement” that joins together *rationality* with *conviction*.<sup>410</sup> In the description of the epistemological skills enumerated in Table 2, the term ‘commitment’ plays a central role in the description of the first four skills; the seventh skill emphasizes the need for a “mental dynamic,” that is, the establishment of an abiding and ongoing way of thinking that invests credence only in what is rationally based. These are the psychological equivalents of the “cement” alluded to in the passage quoted above.

---

<sup>409</sup> Even if and when this happens, the two “bridge problems” are merely shifted to a different level: With neurological region-locating information in hand, merely by virtue of that information it does not follow that we will *therefore* know how to bring about such bridges in individuals who do not have them.

<sup>410</sup> Cf. Bartlett (1969-70).

Although certain of the skills associated with general intelligence may be teachable, and although to some extent some individual students can “be motivated” by their teachers, the kind of strongly internalized mental and psychological commitments to which epistemological intelligence refers are not at present known to be teachable. Perhaps some students can respond to efforts of their teachers who strive to teach epistemological thinking, and can successfully acquire the mental dynamic I am referring to, but generally, at least in my own teaching experience, this is a rarity.

This seems to me much like the fact that only a minority of students of the violin, for example, are able to become highly skilled musicians, while others who may practice just as hard, never can. In music, it is not an intellectual embarrassment to refer to “talent,” an inborn, native capacity to excel as a musician. To say that a high degree of skill reflects “talent” is an admission that we don’t know where it comes from or how to bring it about.

We are much less inclined to accept that in an academic discipline like epistemology anything resembling “talent” should be involved; we *much prefer to believe* that any dedicated student with a college-level amount of general intelligence should be able to become a competent epistemological thinker. At least in connection with the skills I have described, in my experience, this is far from being the case. But more importantly, the strong incompatibilities that exist between Table 2’s skills and Table 1’s psychological limitations lead to the same conclusion.

If epistemological intelligence cannot in any reliable sense be taught, can it then be *learned*? That is, can the relevant skills be learned by an individual whose reasoning capabilities are good and who is, moreover, strongly *self-motivated*? In this case, from what I have observed, the answer is that sometimes, yes, he or she can. But when this happens it is not as though the professor of epistemology has transferred by means of effective communication and training exercises the skills that the “receptive” student is then able to cultivate, internalize, and make the scaffolding of his or her mental operations. In a very real sense, for such a student, the professor is “preaching to the already convinced”—in other words, to the “epistemologically talented.”

...

At one point in the *Divine Comedy*, when Dante meets St. Thomas Aquinas in Paradise, St. Thomas cautions him: “opinion—hasty—often can incline to the wrong side, and then affection for one’s own opinion binds, confines the mind.”<sup>411</sup> —Already in Dante’s 13th century, the closed system of much hu-

---

<sup>411</sup> Mandelbaum & Moser (1980, *Paraiso*, canto XIII, 118-120).

man thinking was evident.

The skills of epistemological intelligence are some of them specialized, and some very general, but taking them together as a group, they make it possible, in the technical philosophical framework of epistemology, for the skilled mind to “think outside the box,” to cut the shackles that hold thought bound to accustomed beliefs that provide gratification, security, and reinforcement of the self. And yet such skills are able to do this in a way that exercises the thinker’s mental restraint in recognizing the limitations of the frames of reference required for knowledge claims to be possible, and by doing this, to avoid transgressing their metalogical horizons. There are some boxes outside of which it is profitable to think, and others outside of which lies only incoherence.

### §10. The projection of transcendence<sup>412</sup>

The tendency has always been strong to believe that whatever receives a name must be an entity of being, having an independent existence of its own. And if no real entity answering to the name could be found, men did not for that reason suppose that none existed, but imagined that it was something peculiarly abstruse and mysterious.

– John Stuart Mill, quoted in Spearman (1927, p. 14)

The human belief in transcendence is a disorder of thought (see {29.5–29.6}): It involves a peculiar variety of conceptual error, an attempted metalogical projection beyond whatever reference frame is in use, plus the predication of independent existence to what is projected. This belief is common among people in all societies. It takes many forms: the child’s belief that a tree must make a crashing sound when there is no one there to hear it; belief in deities who live in a heavenly dimension to which living human beings have no access; belief that a nation possesses a supervening identity and reality for which it makes sense to sacrifice life; the belief that ideologies define an autonomous reality in which the sole, exclusionary, and unique Truth is to be found; and of course the metaphysician’s belief in the reality of objects “in themselves.” When these beliefs are carefully analyzed, none, as we’ve seen in detail in previous chapters, is found to have *possible* meaning, for the meaning they are deceptively thought to have—and the meaning they *can* be

---

<sup>412</sup> Portions of this section are based on Bartlett (2005, pp. 303-304).



thought to have—is a function of and inextricably linked to the reference frames that are necessary to think and articulate those beliefs.

Epistemological contentions like these, however, as we have also seen, can be intellectually slippery, for neither do the opposite claims make sense: It is equally meaningless to claim that the unobserved falling tree does *not* make a sound. —To assert this is not to espouse positivism or operationalism, but rather to endorse the criterion of meaning of referential consistency, which essentially is no more than to recognize that the very identity of objects to which we refer is a framework-relative matter. It is to claim no more than that the point defined by the coordinates (3, -5, 5) has no identity if reference to a three-dimensional Cartesian coordinate system is precluded.

From this point of view, transcendence is a concept, to paraphrase Clifton Fadiman, that looks in vain for a meaning on the scrap heap of popular misuse. We have become so inured to claims that involve its use, and so comforted by the psychological and existential reassurance that it appears to offer, that we cannot see that we hold beliefs in something that is not false, but absurd—that is, incoherent and genuinely meaningless. When we make claims that involve the notion of transcendence, we trespass beyond the boundaries of possible meaningful reference; we step out of bounds, forget where we are, and believe that we have thought a thought that makes good sense, or said something meaningful. —In fact, not only have we not, we cannot *possibly* have said or thought anything meaningful. That very possibility has been undetermined.

For a reflective, epistemologically intelligent person to believe in transcendence is incoherent and intellectual hypocrisy. Epistemological intelligence fuses conviction and rationality in insisting upon a meaningful understanding of reality, one which certainly must, in our terms, be “de-projectively” coherent. Intelligence of this kind is able to free itself from projective beliefs, of which transcendence is unfortunately, as the preceding chapters of this book have made clear, but one among many of the epistemologically delusional beliefs that distort the human comprehension of reality.





## References

The concerns of a metalogic of reference are at once narrow in the sense of specialization of scope, and yet very wide in terms of range of application. There are few books and relatively few articles that specifically have problems of referring, in the wholly abstract and theoretically general sense of this study, as their principal theme. As is made clear elsewhere in this book, the approach to reference and to its metalogic advanced in this study is not linguistically based; it is not philosophy of language. On the one hand, of existing works which in different ways discuss reference in a manner separated from the philosophy of language, none, to my knowledge, attempts to develop a general, theoretical account of the abstract structure of reference with the objective of initiating what we have come to understand as a metalogical investigation of reference and of general problems of reference.

On the other hand, there exist numerous articles and books which offer a rich source for foundational reflections in general, and which provide important tools and ideas useful to the development of the metalogic of reference. The reader who wishes to pursue more specialized lines of thought in the form of applied investigations of particular theoretical frameworks, or the reader who is interested in the most general ramifications of metasytematic reflection, may find certain of the works included in these References helpful as stepping stones to further sources. The list of publications included here, however, is restricted to works referenced in the present study, including, in some instances, related publications. It is not the intention, even if this were a practical possibility, to offer a comprehensive listing of potentially relevant sources.

‘*JSL*’ is used to abbreviate ‘*Journal of Symbolic Logic*’ and ‘*NDJFL*’ to abbreviate ‘*Notre Dame Journal of Formal Logic*’.



Abbott, B. (2000). Presuppositions as nonassertions. *J. of Pragmatics*, 32:10, 1419-1437.

Abbott, B. (2006). Where have some of the presuppositions gone? In Birner, Betty & Ward, Gregory (eds.), *Drawing the boundaries of meaning: Neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*, pp. 1-20. Philadelphia: John Benjamins.

- Abbott, B. (2008). Issues in the semantics and pragmatics of definite descriptions in English. In Gundel, J. K. & Hedberg, N. (eds.), *Reference: Interdisciplinary perspectives*, pp. 61-72. Oxford: Oxford Univ. Press.
- Albritton, Rogers (1957). Present truth and future contingency. *Philosophical Review*, 66, 29-46.
- Anderson, John (1938). The problem of causality. *Australasian J. of Psychology and Philosophy*, XVI, 127-142.
- Anderson, John (1962). *Studies in empirical philosophy*. Sydney, Australia: Angus & Robertson.
- Andreasen, Nancy (1979). Thought, language, and communication disorders, I and II. *Archives of General Psychiatry*, 36, 1315-1321, 1325-1330.
- Angell, R. B. (1964). *Reasoning and logic*. New York: Appleton-Century-Crofts.
- Anscombe, G. E. M (1956). Aristotle and the sea battle. *Mind*, 65, 1-15.
- Åqvist, Lennart (1962). Reflections on the logic of nonsense. *Theoria*, 28:2, 138-157.
- Armstrong, D. M (1986). The nature of possibility. *Canadian J. of Philosophy*, 16:4, 575-594.
- Aschwanden, Christie (2010). Evidence is only part of the story. *Miller-McCune*, May-June, 32-37.
- Asimov, Janet & Asimov, Isaac (1987). *How to enjoy writing: A book to aid and comfort*. New York: Walker.
- Aster, Ernst (1913). *Prinzipien einer Erkenntnislehre*. Leipzig: Quelle & Meyer.
- Austin, J. L. (1962/1955). *How to do things with words*. (William James lectures, 1955.) Cambridge, MA: Harvard Univ. Press.
- Avenarius, Richard Heinrich Ludwig (1907-1908/1888). *Kritik der reinen Erfahrung*. 2 vols. Leipzig: Reisland. 2nd ed. First pub. 1888.
- Ayer, A. J. (1952/1936). *Language, truth, and logic*. New York: Dover. First pub. 1936.
- Bach, Kent (1987). *Thought and reference*. Oxford: Clarendon Press.
- Bachelard, Gaston (1959/1949). The philosophical dialectic of the concepts of relativity. Forrest W. Williams (trans.). In Schilpp (1959/1949), vol. II, pp. 565-580.
- Bachelard, Suzanne (1957). *La logique de Husserl: Étude sur "Logique formelle et logique transcendantale."* Paris: Presses Universitaires de France.

- Bardon, A. (2005). Performative transcendental arguments. *Philosophia*, 33, 69-95.
- Bartlett, Steven James & Suber, Peter (eds.) (1987). *Self-reference: Reflections on reflexivity*. Dordrecht, Holland: Martinus Nijhoff (now re-issued by Springer Verlag).
- Bartlett, Steven James (1969-70). Conviction and rationality. A paper presented to the Fellows, Center for the Study of Democratic Institutions, Santa Barbara, CA, during the academic year 1969-1970. Available online from PhilPapers and Willamette University.
- Bartlett, Steven James (1970). *A relativistic theory of phenomenological constitution: A self-referential, transcendental approach to conceptual pathology*. vol. I: French; vol. II: English. Ph.D. dissertation, Université de Paris. University Microfilms Intl. #7905583. Published online by the Erasmus University of Rotterdam; also available online from PhilPapers.
- Bartlett, Steven James (1973). *Validity: A learning game approach to mathematical logic*. Hartford, CT: Lebon Press. Rev. ed. 1974. Available online from PhilPapers and Willamette University.
- Bartlett, Steven James (1974). Fenomenologia tego, co implikowane. *Roczniki Filozoficzne*, XXII:1, 73-89. In Polish. For English, see Bartlett (1975a). Available online from PhilPapers.
- Bartlett, Steven James (1975). *Metalogic of reference: A study in the foundations of possibility*. Starnberg, Germany: Max-Planck-Gesellschaft.
- Bartlett, Steven James (1975a). Phenomenology of the implicit. *Dialectica*, 29:2-3. Published earlier in Polish translation, Bartlett (1974). Available online from PhilPapers.
- Bartlett, Steven James (1975b). Towards a unified concept of reality. *ETC.: A Review of General Semantics*, 32:1, 43-49. Available online from PhilPapers.
- Bartlett, Steven James (1976). The idea of a metalogic of reference. *Methodology & Science* 9:3, pp. 85-92. Available online from PhilPapers.
- Bartlett, Steven James (1976-77). *Patterns of problem-solving: A final report of work undertaken with the support of the Lilly Endowment*. Saint Louis, MO: Saint Louis Univ. Available online from Willamette University.
- Bartlett, Steven James (1978-79). Cognitive skills in philosophy: A teacher's guide. *Aitia*, 6:3, 12-21. Available online from PhilPapers.

- Bartlett, Steven James (1978a). A metatheoretical basis for interpretations of problem-solving behavior. *Methodology and Science*, 11:2, 59-85. Available online from PhilPapers and Willamette University.
- Bartlett, Steven James (1978b). The use of protocol analysis in philosophy. *Metaphilosophy*, 9:3-4, 324-336. Available online from PhilPapers.
- Bartlett, Steven James (1979). Review of Boyle, Grisez, & Tollefsen (1976). *Review of Metaphysics*, 32:4, 738-740. Available online from PhilPapers.
- Bartlett, Steven James (1980). Self-reference, phenomenology, and philosophy of science. *Methodology and Science: Interdisciplinary J. for the Empirical Study of the Foundations of Science and Their Methodology*, 13:3, pp. 143-167. Available online from PhilPapers and Willamette University.
- Bartlett, Steven James (1982). Referential consistency as a criterion of meaning. *Synthese*, 52, 267-82. Available online from PhilPapers and Willamette University.
- Bartlett, Steven James (1983a). *Conceptual therapy: An introduction to framework-relative epistemology*. St. Louis, MO: Studies in Theory and Behavior. Available online from Erasmus University Rotterdam, Willamette University, and PhilPapers.
- Bartlett, Steven James (1983b). Philosophy as conceptual therapy. *Educational Resources Information Center*, Natl. Inst. of Education, Dept. of Health, Education, & Welfare, May. Document #ED 224 402. Available online from PhilPapers and Willamette University.
- Bartlett, Steven James (1986a). Philosophy as ideology. *Metaphilosophy*, 17:1, 1-13. Available online from PhilPapers.
- Bartlett, Steven James (1986b). Narcissism and philosophy. *Methodology and Science: Interdisciplinary J. for the Empirical Study of the Foundations of Science and Their Methodology*, 19:1, 16-26. Available online from PhilPapers.
- Bartlett, Steven James (1987). Varieties of self-reference. In Bartlett & Suber (1987), pp. 5-28. Available online from Willamette University and PhilPapers.
- Bartlett, Steven James (1988). Hoisted by their own petards: Philosophical positions that self-destruct. *Argumentation*, 2, 221-232. Available online from PhilPapers.
- Bartlett, Steven James (1989). Psychological underpinnings of philosophy. *Metaphilosophy*, 20:3 & 4, 295-305. Available online from PhilPapers.

- Bartlett, Steven James (1992a). *Reflexivity: A source book in self-reference*. Elsevier Science Publishers. Available online from PhilPapers, the Erasmus University of Rotterdam, and Willamette University.
- Bartlett, Steven James (1992b). The role of reflexivity in understanding human understanding. In Bartlett (1992a), pp. 3-18. Available online from PhilPapers.
- Bartlett, Steven James (2002). Roots of human resistance to animal rights: Psychological and conceptual blocks. *Animal Law*, 8, 143-76. Available online from *Animal Law* and PhilPapers.
- Bartlett, Steven James (2005). *The pathology of man: A study of human evil*. Springfield, IL: Charles C. Thomas.
- Bartlett, Steven James (2008). The humanistic psychology of human evil: Ernest Becker and Arthur Koestler. *J. of Humanistic Psychology*, 48:3, 340-363.
- Bartlett, Steven James (2011). *Normality does not equal mental health: The need to look elsewhere for standards of good psychological health*. Santa Barbara, CA: Praeger.
- Bartlett, Steven James (2013). The dilemma of abnormality. In Plante, Thomas G. (ed.), *Abnormal psychology across the ages*, vol. 3, Chapter 1, pp. 1-20. Santa Barbara, CA: Praeger.
- Bartlett, Steven James (2015). *The species problem and its logic: Inescapable ambiguity and framework relativity*. Published online Fall, 2015 by arXiv.org, in England by CogPrints, in France by the Centre pour la Communication Scientifique Directe's HAL, and in the U.S. by PhilSci.
- Bartlett, Steven James (2016). *Paratheism: A proof that God neither exists nor does not exist*. Published online; available from PhilPapers and Willamette University.
- Bartlett, Steven James (2017a). Encounters with intellectual suppression. Published online; available from bmartin.cc.
- Bartlett, Steven James (2017b). Peer review—an insult to the reader and to society: Milton's view. Published online; available from PhilPapers and CogPrints.
- Bartlett, Steven James (2019). A code of conduct for peer reviewers and editors. Published online; available from PhilPapers and Social Science Research Network (SSRN).
- Bateson, Gregory & Jackson, Don D. (1964). Social factors and disorders of communication: Some varieties of pathogenic organization. *Research Publi-*

- cations: Association for Research in Nervous and Mental Disease*, 42, 270-290.
- Bateson, Gregory (1972). *Steps to an ecology of mind*. New York: Ballantine Books.
- Baylis, Charles A. (1936). Are some propositions neither true nor false? *Philosophy of Science*, 3, 156-166.
- Beall, J.C. (2000). Fitch's proof, verificationism, and the knower paradox. *Australasian J. of Philosophy*, 78, 241-247.
- Beaney, M., (ed.) (1997). *The Frege reader*. Oxford: Basil Blackwell.
- Beaver, David & Kraemer, Emiel (2001). A partial account of presupposition projection. *J. of Logic, Language and Information*, 10, 147-182.
- Beaver, David (1992). The kinematics of presupposition. In Dekker, P. & Stockhof, M. (eds.), *Proceedings of the Eighth Amsterdam Colloquium*, Univ. of Amsterdam: ILLC, 17-36.
- Beaver, David (1997). Presupposition. In Benthem, J. van & Meulen, A. ter (eds.), *The Handbook of Logic and Language*, pp. 939-1008. Amsterdam: Elsevier.
- Beaver, David (2001). *Presupposition and assertion in dynamic semantics*. Stanford, CA: CSLI Publications.
- Beaver, David (2008). As brief as possible (but no briefer). *Theoretical Linguistics*, 34, 213-228.
- Becker, Ernest (1975). *Escape from evil*. New York: Free Press.
- Beeley, Philip (2003). Leibniz on the limits of human knowledge: With a critical edition of *Sur la calculabilité du nombre de toutes les connaissances possibles* and English translation. *Leibniz Review*, 13, 83-91. [Beeley's translation appears on pp. 99-101.]
- Benacerraf, Paul & Putnam, Hilary (eds.) (1964). *Philosophy of mathematics: Selected readings*. Englewood Cliffs, NJ: Prentice-Hall.
- Bentham, Jeremy (1876). *An introduction to the principles of morals and legislation*. Oxford: Clarendon Press.
- Berkeley, George (1999/1710-1713). Howard Robinson (ed.). *Principles of human knowledge and Three dialogues*. Oxford: Oxford Univ. Press. [*Principles* first pub. 1710; *Dialogues* first pub. 1713.]
- Berto, Francesco (2007). *How to sell a contradiction: The logic and metaphysics of inconsistency*. Strand, London: College Publications.



- Bevington, Philip R., & Robinson, D. Keith (2010/1969). *Data reduction and error analysis for the physical sciences*. 3rd ed. Boston: McGraw-Hill. First pub. 1969.
- Bieri, P.; Horstman, R.-P.; & Krüger, L. (eds.) (1979). *Transcendental arguments and science*. Dordrecht, Holland: D. Reidel.
- Binet, Alfred & Henri, Victor (1895). La psychologie individuelle. *L'Année psychologique*, 2, 411-465.
- Black, Max (1959/1933). *The nature of mathematics*. Paterson, NJ: Littlefield, Adams & Co. First pub. 1933.
- Bloomfield, L. (1933). *Language*. New York : Holt, Rinehart, and Winston.
- Bochvar, D. A. (1939). Ob odnom trékhznačnom isčislénii i égo priménénii k analizu paradoksov klassičeskogo rassírennogo funkcional'nogo isčisléniiá [On a three-valued logical calculus and its application to the analysis of contradictions]. *Matématiceskij sbornik*, 4, 287-308.
- Bochvar, D. A. (1943). K voprosu o néprotivoréčivosti odnogo trékhznačnogo isčisléniiá [On the consistency of a three-valued calculus]. *Matématiceskij sbornik*, 12, 353-369.
- Bohm, David (1952). A suggested interpretation of the quantum theory in terms of hidden variables. *Physical Review*, LXXXV: Part I, 166-179; Part II, 180-193.
- Bohm, David (1962). A proposed topological formulation of the quantum theory. In Good (1962), pp. 302-314.
- Bohr, Niels (1934). *Atomic theory and the description of nature*. Cambridge: Cambridge Univ. Press.
- Bohr, Niels (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 48, 696-702.
- Bohr, Niels (1958). *Atomic physics and human knowledge*. New York: John Wiley.
- Bohr, Niels (1959/1949). Discussion with Einstein on epistemological problems in atomic physics. In Schilpp (1959/1949), vol. 1, pp. 199-241. New York: Harper and Row. First printed 1949.
- Bohr, Niels (1996). *Foundations of quantum physics II (1933-1958)*. Jørgen Kalckar (ed.). Niels Bohr, *Collected Works*, vol. 4. Amsterdam: North-Holland.
- Bolzano, Bernard (2004/1851). *Paradoxes of the infinite*. In Bolzano, Bernard, *The mathematical works of Bernard Bolzano*, (Steve Russ, ed.), pp. 591-678. Oxford: Oxford Univ. First pub. 1851.

- Borges, Jorge Luis (1998/1941). *Collected fictions*. Andrew Hurley (trans.). New York: Penguin. Cited passage first pub. 1941.
- Bourget, D. & Chalmers, D. J. (2013). What do philosophers believe? *Philosophical Studies*, 3, 1–36. Also available from PhilPapers.
- Bouwisma, O. K. (1965). Descartes' evil genius. In Sesonske and Fleming (1965), pp. 26-36.
- Boyle, Joseph M., Jr.; Grisez, Germain; & Tollefsen, Olaf (1972). Determinism, freedom, and self-referential arguments. *Rev. of Metaphysics*, XXVI, 3-37.
- Boyle, Joseph M., Jr.; Grisez, Germain; & Tollefsen, Olaf (1976). *Free choice: A self-referential argument*. Notre Dame, IN: Univ. of Notre Dame Press.
- Bredeck, Elizabeth (1992). *Metaphors of knowledge: Language and thought in Mauthner's Critique*. Detroit: Wayne State Univ. Press.
- Brentano, Franz (1924-25/1874). *Psychologie vom empirischen Standpunkte*. Leipzig: Felix Meiner. First pub. 1874.
- Bridgman, P. W. (1936). *The nature of physical theory*. New York: Dover.
- Bridgman, P. W. (1959). *The way things are*. Cambridge, MA: Harvard Univ. Press.
- Bridgman, P. W. (1961/1927). *The logic of modern physics*. New York: Macmillan. First pub. 1927.
- Brillouin, Louis Marcel. (1970). *Relativity reexamined*. New York: Academic Press.
- Brouwer, L. E. J. (1964/1912). Intuitionism and formalism. Arnold Dresden (trans.). In Benacerraf & Putnam (1964), pp. 66-77.
- Brouwer, L. E. J. (1964/1940). Consciousness, philosophy, and mathematics. In Benacerraf & Putnam (1964), pp. 78-84.
- Bruner, J. S. & Tagiuri, R. (1954). The perception of people. In Lindzey, G. (ed.), *Handbook of social psychology*, pp. 634-54. Reading, MA: Addison-Wesley.
- Bruner, J. S. (1961). The act of discovery. *Harvard Educ. Rev.*, 31, 21-32.
- Buchanan, Scott (1927). *Possibility*. London: Kegan, Paul, Trench, Trubner.
- Buckley, Walter (1972). A systems approach to epistemology. In Klir, George J. (ed.), *Trends in general systems theory*, pp. 188-202. New York: John Wiley.
- Bull, R. A. (1964). An axiomatization of Prior's modal calculus Q. *NDJFL*, 5, 211-214.
- Burt, Cyril (1940). *The factors of the mind: An introduction to factor-analysis in psychology*. London: Univ. of London Press.

- Butler, R. J. (1955). Aristotle's sea fight and three-valued logic. *Philosophical Review*, 54, 264-274.
- Calder, Allan (1980). Letters. *Scientific American*, 242:2, 6.
- Carnap, Rudolf (1934/1931). *La science et la métaphysique devant l'analyse logique du langage*. Ernest Vuillemin (trans.). Paris: Hermann. First pub. 1931.
- Carnap, Rudolf (1935). *Philosophy and logical syntax*. London: Kegan Paul.
- Carnap, Rudolf (1942). *An introduction to semantics*. Cambridge, MA: Harvard Univ. Press.
- Carnap, Rudolf (1950). Empiricism, semantics, and ontology. *Revue Internationale de Philosophie*, 4, 20-40. [Reprinted in Carnap (1956/1947), pp. 205-221; references are to this edition. Also reprinted in Benacerraf & Putnam (1964), pp. 234-242.]
- Carnap, Rudolf (1956/1947). *Meaning and necessity: A study in semantics and modal logic*. Chicago: Univ. of Chicago Press. First pub. 1947.
- Carnap, Rudolf (1959/1932). The elimination of metaphysics through logical analysis of language. Arthur Pap (trans.). In Ayer, A. J. (ed.), *Logical positivism*, pp. 60-81. New York: Free Press. First pub. 1932.
- Carnap, Rudolf (1963). Intellectual autobiography. In Schilpp, Paul Arthur (ed.), *The philosophy of Rudolf Carnap*, pp. 1-84. La Salle, IL: Open Court.
- Carnap, Rudolf (1967/1928). *The logical structure of the world and Pseudoproblems in philosophy*. Rolf A. George (trans.). Berkeley, CA: Univ. of Calif. Press. Originally published as *Der logische Aufbau der Welt* and as *Scheinprobleme in der Philosophie: Das Fremdpsychische und der Realismusstreit*. Berlin/Schlachtensee: Weltkreis-Verlag, 1928.
- Castagnoli, L. (2007). Everything is true, everything is false: Self-refutation arguments from Democritus to Augustine. *Antiquorum philosophia*, 1, 11-74.
- Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge: Cambridge Univ. Press.
- Christodoulou, G. N. (1991). The delusional misidentification syndromes. *Brit. J. of Psychiatry*, 159: Suppl. 14, 65-9.
- Church, Alonzo (1939-40). Review of D. A. Bochvar's "On a three-valued logical calculus and its application to the analysis of contradictions." *JSL*, 4, 1939, 98-99; correction *JSL*, 5, 1940, 119.
- Church, Alonzo (2009). Referee reports on Fitch's "A Definition of Value." In Salerno (2009), pp. 13-20.

- Clifford, William K. [Kingdon] (1877). The ethics of belief. *Contemporary Review*, 29, 289-309. Reprinted in William K. Clifford, *Lectures and Essays*. Leslie Stephen & Frederick Pollock (eds.). London: Macmillan, 1886. Currently in print in William K. Clifford, *The ethics of belief and other essays*. New York: Prometheus Books, 1999. Available online.
- Clifford, William K. [Kingdon] (1878). On the nature of things-in-themselves. *Mind*, 3:9, 57-67.
- Cocchiarella, Nino B. (1966). A logic of actual and possible objects. *JSL*, 31, 688-689.
- Coleridge, Samuel Taylor (1817). *Biographia literaria; or biographical sketches of my literary life and opinions*. 2 vols. London: Rest Fenner.
- Collingwood, R. C. (1933). *An essay on philosophical method*. London: Oxford Univ. Press.
- Collingwood, R. G. (1940). *An essay on metaphysics*. Oxford: Clarendon Press.
- Coltheart, Max & Davies, Martin (eds.) (2000). *Pathologies of belief*. Malden, MA: Blackwell.
- Comey, David Dinsmore (1965). Review of V.A. Smirov's "The logical views of N. A. Vasil'ev." *JSL*, 30, 368-370.
- Coreth, Emerich (1961). *Metaphysik. Eine methodisch-systematische Grundlegung*. Innsbruck-Vienna-Munich: Tyrolia-Verlag.
- Crawford, Patricia (1961). Kant's theory of philosophical proof. *Kant-Studien* 53:3, 257-268.
- Curry, Haskell B. (1957). *Outlines of a formalist philosophy of mathematics*. Amsterdam: North-Holland.
- Danto, Arthur C. (1971/1968). *What philosophy is: A guide to the elements*. New York: Harper. First pub. 1968.
- Davidson, Donald & Harman, Gilbert (eds.) (1972). *Semantics of natural language*. New York: Humanities Press.
- Davis, W. (2002). *Meaning, expression, and thought*. New York: Cambridge Univ. Press.
- Dawkins, Richard (2006). *The God delusion*. New York: Houghton Mifflin.
- de Broglie, Louis (1926). *Ondes et mouvements*. Paris: Gauthier-Villars.
- de Broglie, Louis (1964). *The thermodynamics of the isolated particle (or the hidden thermodynamics of particles)*. D. H. Delphenich (trans.). Paris: Gauthier-Villars. (Edited lectures from 1961-1962.)

- Dilman, Ilham (1996). *Wisdom. Philosophy*, 71:278, 577-90.
- Dilman, Ilham (ed.) (1984). *Philosophy and life: Essays on John Wisdom*. Dordrecht: Nijhoff.
- Donceel, Joseph (1974). Transcendental Thomism. *Monist*, 58, 67-85.
- Dreyfus, Hubert L. (1972). *What computers can't do: A critique of artificial intelligence*. New York: Harper and Row.
- Driesch, Hans (1923/1912). *Ordnungslehre*. Jena: Eugen Diederichs. First pub. 1912.
- Drieschner, Michael (1970). *Quantum mechanics as a general theory of objective prediction*. Doctoral dissertation, Universität Hamburg.
- Drieschner, Michael (1974). Objekte der Naturwissenschaft. *Neue Hefte für Philosophie*, 6:7, 104-128.
- Edgington, D. (2010). Possible knowledge of unknown truth. *Synthese*, 173, 41-52.
- Einstein, Albert (1905). Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt [On a heuristic point of view concerning the production and transformation of light]. *Annalen der Physik*, 17, 132-148.
- Einstein, Albert (1916). Ernst Mach [obituary]. *Physikalische Zeitschrift*, 17:7, 101-104.
- Einstein, Albert (1916a). Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, IV, 49:7, 769-822. Pub. separately, Leipzig: Barth, 1916. Pub. 2009 in English, Robert W. Lawson (trans.). *Relativity: Einstein's theory of spacetime, time dilation, gravity, and cosmology*. St. Petersburg, FL: Red and Black Publishers.
- Einstein, Albert (1920/1917a). *Relativity, the special and the general theory: A popular exposition*. Robert W. Lawson (trans.). London: Methuen. First pub. in German, 1917.
- Einstein, Albert (1920/1917b). *Über die spezielle und die allgemeine Relativitätstheorie, Gemeinverständlich*. 5. Aufl. Braunschweig: Friedr. Vieweg & Sohn. First pub. 1917.
- Einstein, Albert (1921). A brief outline of the development of the theory of relativity. Robert William Lawson (trans.). *Nature*, 106:2677, 782-784.
- Einstein, Albert (1922). *The meaning of relativity*. Edwin Plimpton Adams (trans.). Princeton, NJ: Princeton Univ. Press.

- Einstein, Albert (1928). Review by Albert Einstein of Émile Meyerson's *La Déduction relativiste*. In Meyerson (1985/1925), Appendix I, pp. 252-256, André Metz (trans.). [Based on the original publication, "À propos de *La Déduction relativiste* de M. Émile Meyerson," *Revue philosophique de la France et de l'étranger*, 105, 1928, 161-166, and also based on the unpublished German text in the Einstein archive.]
- Einstein, Albert (1931). Maxwell's influence on the development of the conception of physical reality. In Thomson, J. J., *James Clerk Maxwell: A commemoration volume, 1831-1931*, pp. 66-73. Cambridge: Cambridge Univ. Press.
- Einstein, Albert (1934). *The world as I see it*. New York: Covici-Friede.
- Einstein, Albert (1934/1933). On the method of theoretical physics. *Philosophy of Science*, 1:2, 162-169. Oxford: Clarendon Press. First pub. 1933.
- Einstein, Albert (1936). Physik und Realität. *J. of the Franklin Inst.*, 221, 313-347.
- Einstein, Albert (1944). Remarks on Russell's theory of knowledge. In Schilpp, Paul Arthur (ed.), *The philosophy of Bertrand Russell*, pp. 277-291. Evanston: Northwestern Univ. Press.
- Einstein, Albert (1950). *Out of my later years*. New York: Philosophical Library.
- Einstein, Albert (1959/1949a). Remarks concerning the essays brought together in this co-operative volume. In Schilpp (1959/1949), vol. II, pp. 665-688.
- Einstein, Albert (1959/1949b). Autobiographical notes. In Schilpp (1959/1949), vol. I, pp. 3-95.
- Einstein, Albert (1965/1956). La relativité et le problème de l'espace." In *La relativité: La théorie de la relativité restreinte et générale*, Quatrième Partie, pp. 155-179, written by Einstein for this French edition. Maurice Solovine (trans.). Paris: Petite Bibliothèque Payot. First pub. in French 1956.
- Einstein, Albert (2006/1910). On Boltzmann's Principle and some immediate consequences thereof. B. Duplantier & E. Parks (trans.). Unpublished manuscript of a 1910 lecture by Einstein. In Damour, T. (ed.), *Einstein, 1905-2005, Poincaré Seminar 2005*, pp. 183-199. Boston: Birkhäuser Verlag.
- Einstein, Albert; Podolsky, Boris; & Rosen, Nathan (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, ser. 2, 47, 777-780.
- Erwin, Edward (1970). *The concept of meaninglessness*. Baltimore, MD: Johns Hopkins Press.

- Evans G., & McDowell, J., (eds.) (1976). *Truth and meaning: Essays in semantics*. Oxford: Oxford Univ. Press.
- Fara, Michael (2010). Knowability and the capacity to know. *Synthese*, 173, 53-73.
- Farrell, Brian A. (1946). An appraisal of therapeutic positivism. *Mind*, 55, 25-48 (Part I), 133-50 (Part II).
- Farrell, Brian A. (1990). The meaning of life: A psycho-logical exploration. In Roux, A. P. J. (ed.), *Life, world and meaning: Essays in honor of Professor Michael Macnamara*, pp. 1-8. Pretoria, S. Africa: Univ. of South Africa.
- Feng, Justin C. (2017). *A poor man's introduction to tensors*. Austin, TX: Dept. of Physics.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row Peterson.
- Festinger, L.; Riecken, H.W.; & Schachter, S. (1956). *Prophecy fails*. Minneapolis, MN: Univ. of Minnesota Press.
- Feys, F. & Fitch, F. B. (1969). *Dictionary of mathematical logic*. Amsterdam: North-Holland.
- Findlay, J. N. (1963). *Meinong's theory of objects and values*. Oxford: Clarendon Press.
- Finnis, John (2004). Self-referential (or performative) inconsistency. *Proceedings of the American Catholic Philosophical Assoc.*, 78, 13-22.
- Fitch, F. B. (1942). A basic logic. *JSL*, 7, 105-114.
- Fitch, F. B. (1944). Representations of calculi. *JSL*, 9, 57-62.
- Fitch, F. B. (1944a). A minimum calculus for logic. *JSL*, 9, 89-94.
- Fitch, F. B. (1946). Self-reference in philosophy. *Mind*, 55, 64-73.
- Fitch, F. B. (1948). An extension of basic logic. *JSL*, 13, 95-106.
- Fitch, F. B. (1949). The Heine-Borel Theorem in extended basic logic. *JSL*, 14, 9-15.
- Fitch, F. B. (1950-51). A demonstrably consistent mathematics – Part I: *JSL*, 15, 17-24; Part II: *JSL*, 16, 121-124.
- Fitch, F. B. (1950a). A further consistent extension of basic logic. *JSL*, 14, 209-217.
- Fitch, F. B. (1950b). Actuality, possibility, and being. *Rev. of Metaphysics*, 3:3, 367-384.
- Fitch, F. B. (1952). *Symbolic logic: An introduction*. New York: Ronald Press.

- Fitch, F. B. (1963). A logical analysis of some value concepts. *JSL*, 28, 135-142. Reprinted in Salerno (2009), pp. 21–28.
- Fitch, F. B. (1963-64). Universal metalanguages for philosophy. *Review of Metaphysics*, 17, 396-402.
- Flügel, John Carl (1951/1933). *A hundred years of psychology: 1833-1933; with additional part on developments 1933-1947*. London: G. Duckworth. First pub. 1933.
- Frege, Friedrich Ludwig Gottlob (1964/1884). The concept of number. Michael S. Mahoney (trans.). In Benacerraf & Putnam (1964), pp. 85-112. First pub. 1884.
- Frege, Friedrich Ludwig Gottlob (1964/1893-1903). *The basic laws of arithmetic, Exposition of the system*. Montgomery Furth (trans. & ed.). Berkeley, CA: Univ. of California Press. (Originally published as *Grundgesetze der Arithmetik*, 2 Bde., 1893–1903.)
- Frege, Friedrich Ludwig Gottlob (1979). *Posthumous writings*. London: Blackwell.
- Freire Olival, Jr. (2005). Science and exile: David Bohm, the hot times of the Cold War, and his struggle for a new interpretation of quantum mechanics. *Historical Studies in the Physical and Biological Sciences*, 36:1, 1–34.
- Freud, Sigmund (1930/1952). *Civilization and its discontents*. J. Riviere (trans.). *Great books of the western world*, vol. 54, 767-802. Chicago: Encyclopaedia Britannica. First pub. 1930.
- Freud, Sigmund (1953-74). *The standard edition of the complete psychological works of Sigmund Freud*. 24 vols. J. Strachey (trans. & ed.); A. Freud; A. Strachey; & A. Tyson (trans.). London: Hogarth.
- Freud, Sigmund (1953-74/1894). *The neuro-psychoses of defense*. In Freud (1953-74), vol. 3, pp. 43-69. First pub. 1894.
- Freud, Sigmund (1953-74/1896). *Further remarks on the neuro-psychoses of defense*. In Freud (1953-74), vol. 3, pp. 159-188. First pub. 1896.
- Freud, Sigmund (1962/1912-13). *Totem and taboo: Some points of agreement between the mental lives of savages and neurotics*. J. Strachey (trans.). New York: W. W. Norton. First pub. 1912-13.
- Furberg, Mats (1973). On questions and pseudo-problems. In *Modality, morality and other problems of sense and nonsense: Essays dedicated to Sören Halldén* (no editor indicated), pp. 85-96. Lund, Sweden: CWK Gleerup Bokförlag.



- Gagné, R. M. (1966). Human problem-solving: Internal and external events. In Kleinmuntz, B. (ed.), *Problem solving – Research, method, and theory*, pp. 128-148. New York: Wiley.
- Gal, David & Rucker, Derek D. (2010). When in doubt, shout! Paradoxical influences of doubt on proselytizing. *Psychological Science*, 21:11, 1701-1707.
- Gardner, Howard (1993). *Multiple intelligences: The theory in practice*. New York: BasicBooks.
- Gardner, Howard (1993/1983). *Frames of mind: The theory of multiple intelligences*. New York: BasicBooks. First pub. 1983.
- Gauss, Johann Karl Friedrich (1828). *Disquisitiones generales circa superficies curvas*. Göttingen: Typis Dieterichianis.
- Getzels, J. W. & Jackson, P. W. (1962). *Creativity and intelligence*. New York: Wiley.
- Ghirardi, G.C.; Rimini, A.; & Weber, T. (1986). Unified dynamics for microscopic and macroscopic systems. *Physical Review D (Particles and Fields)*, 34:2, 470-491.
- Ghiselin, Brewster (1985/1952). *The creative process: A symposium*. Berkeley, CA: Univ. of California Press. First pub. 1952.
- Ginsberg, Mitchell (1972). The entailment-presupposition relationship. *NDJFL*, XIII:4, 511-515.
- Goddard, L. (1964). Sense and nonsense. *Mind*, LXXIII:291, 309-331.
- Goddard, L. (1966). Predicates, relations, and categories. *Australasian J. of Philosophy*, 44, 139-171.
- Goddard, Leonard & Routley, Richard (1973). *The logic of significance and context*. Vol. 1. Edinburgh: Scottish Academic Press.
- Goethe, Johann Wolfgang von (1969/1800). Natur und Kunst. In Trunz, Erich (ed.), *Goethes Werke: Hamburger Ausgabe*, vol. I, p. 245. München: Beck.
- Gold, Ian & Hohwy, Jakob (2000). Rationality and schizophrenic delusion. In Coltheart & Davies (2000), pp. 145-165.
- Goleman, Daniel (1995). *Emotional intelligence*. New York: Bantam.
- Good, Irving John (ed.) (1962). *The scientist speculates: An anthology of partly-baked ideas*. London: Heinemann.
- Goodman, Nelson (1951). *The structure of appearance*. Cambridge, Mass.: Harvard Univ. Press. [Rewritten from *A study of qualities*, doctoral thesis, Harvard University, 1940.]

- Goodman, Nelson (1964). The world of individuals. In Benacerraf & Putnam (1964), pp. 197-209.
- Gottlieb, Dale (1978). No entity without identity. In Shahan, Robert W. & Swoyer, Chris (eds.), *Essays on the philosophy of W. V. Quine*, pp. 79-96. Norman, OK: Univ. of Oklahoma Press.
- Graffin, Gregoriy W. & Provine, William B. (2007). Evolution, religion, and free will. *American Scientist*, 95, July-August, 294-97.
- Grier, Michelle (2001). *Kant's doctrine of transcendental illusion*. Cambridge: Cambridge Univ. Press.
- Griffiths, A. Phillips (1969). Transcendental arguments. *Proceedings of the Aristotelian Society*, Supplementary Volumes, 43, 165-180.
- Gruber, H. E.; Terrell, G.; & Wertheimer, M. (eds.) (1962). *Contemporary approaches to creative thinking*. New York: Atherton Press.
- Grynepas, Jérôme (1961). Remarques sur le problème de la critique philosophique en philosophie. *Revue internationale de philosophie*, XV:58, 410-432.
- Guilford, J. P. (1958). Basic traits in intellectual performance. In Taylor, C. W. (ed.), *The second University of Utah research conference on the identification of scientific talent*. Salt Lake City: Univ. of Utah Press.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Gurwitsch, Aron (1966). An apparent paradox in Leibnizianism. *Social Research*, 33:1, 47-64.
- Hacker, P. M. S. (2009). Philosophy: A contribution to human understanding, not to human knowledge. *Royal Institute of Philosophy Supplement*, 65, 129-153.
- Hadamard, Jacques (1954/1945). *An essay on the psychology of invention in the mathematical field*. New York: Dover. First pub. 1945.
- Hájek, P. & Pudlák, P. (1993). *Metamathematics of first-order arithmetic*. Berlin: Springer.
- Halldén, Sören (1949). *The logic of nonsense*. Uppsala, Sweden: Appelbergs Boktryckeri.
- Halldén, Sören (1967/1961). *Universem, döden och den logiska analysen* [Universe, death and logical analysis]. Lund: Gleerups. Orig. pub. 1961.
- Halmos, Paul (1951). *Introduction to Hilbert space and the theory of spectral multiplicity*. New York: Chelsea Publishing Co.
- Hancock, Roger (1960). Presuppositions. *Philosophical Quarterly*, 10:38, 73-78.

- Hanson, Norwood Russell (1958). *Patterns of discovery*. Cambridge: University Press.
- Hanson, Norwood Russell (1971). *What I do not believe and other essays*. Dordrecht-Holland: D. Reidel.
- Hardin, C. L. & L. Maffi (eds.) (1997), *Color categories in thought and language*. Cambridge: Cambridge Univ. Press.
- Harris, Sam (2012). *Free will*. New York: Free Press.
- Hartmann, Klaus (1966). On taking the transcendental turn. *Rev. of Metaphysics*, 20:2, 223-249.
- Heelan, Patrick A. (1965). *Quantum mechanics and objectivity: A study of the physical philosophy of Werner Heisenberg*. Haag: M. Nijhoff.
- Heelan, Patrick A. (1967). Horizon, objectivity and reality in the physical sciences. *International Philosophical Quarterly*, 7, 375-412.
- Heffernan, Margaret (2011). *Willful blindness: Why we ignore the obvious at our peril*. New York: Walker.
- Hegel, Georg Wilhelm Friedrich (1969/1812). *Hegel's science of logic*. A. V. Miller (trans.). London: Allen and Unwin.
- Heidegger, Martin (1962/1927). *Being and Time*. John Macquarrie & Edward S. Robinson (trans.). New York: Harper. First pub. 1927.
- Heisenberg, Werner (1927), Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43:3-4, 172-198.
- Heisenberg, Werner (1930). *The physical principles of the quantum theory*. Carl Eckart & Frank C. Hoyt (trans.). New York: Dover. First German pub. also 1930.
- Heisenberg, Werner (1955). The development of the interpretation of the quantum theory. In Pauli, Wolfgang (ed.), *Niels Bohr and the development of physics*, pp. 12-29. London: Pergamon.
- Heisenberg, Werner (1958). *Physics and philosophy: The revolution in modern science*. New York: Harper & Row.
- Heisenberg, Werner (1969). *Der Teil und das Ganze: Gespräche im Umkreis der Atomphysik*. München: R. Piper.
- Henderson, G. P. (1959). Is there only one correct system of modal logic? Part II. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 33, 41-56.
- Hesse, Hermann (1970/1943). *Magister Ludi (The Glass Bead Game)*. Richard & Clara Winston (trans.). Holt, Rinehart and Winston (Bantam). First pub. 1943.

- Heyting, Arend (1964). The intuitionist foundations of mathematics. In Benacerraf & Putnam (1964), pp. 42-54.
- Heyting, Arend (1964/1936). Disputation. In Benacerraf & Putnam (1964), pp. 55-65. First pub. 1936.
- Hilbert, David (1964/1925). On the infinite. Erna Putnam & Gerald J. Massey (trans.). In Benacerraf & Putnam (1964), pp. 134-151.
- Hintikka, Jaakko (1957). Modality as referential multiplicity. *Ajatus*, 20, 49-64.
- Hintikka, Jaakko (1957a). Necessity, universality, and time in Aristotle. *Ajatus*, 20, 65-90.
- Hintikka, Jaakko (1962). *Knowledge and belief: An introduction to the logic of the two notions*. Ithaca, N.Y.: Cornell Univ. Press.
- Hintikka, Jaakko (1965/1962). *Cogito, ergo sum: Inference or performance*. In Sesonke, Alexander & Fleming, Noel (eds.), *Meta-meditations: Studies in Descartes*, pp. 50-76. Belmont, CA: Wadsworth. First pub. 1962.
- Hjelmslev, Louis (1947). Structural analysis of language. *Studia Linguistica: Revue de linguistique générale et comparée*, vol. 1, pp. 69-78.
- Hocking, W. E. (1939). *Types of philosophy*. New York: Charles Scribner's Sons.
- Hocutt, Max O. (1972). Is epistemic logic possible? *NDJFL*, XIII:4, 433-453.
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Rev. of Educational Research*, 67, 88-140.
- Holton, Gerald (2006). Einstein and the shaping of our imagination. In Goldsmith, Donald & Bartusiak, Marcia (eds.), *E = Einstein: His life, his thought, and his influence on our culture*, pp. 263-285. New York: Sterling Publishing.
- Horwich, P. (2005). *Reflections on meaning*. Oxford: Clarendon Press.
- Hueffer, Ford M. [Ford, Ford Madox] and Conrad, Joseph (1901). *The inheritors: An extravagant story*. New York: McClure, Phillips & Co.
- Hume, David (1888). *A treatise of human nature*. L. A. Selby-Bigge (ed.). Oxford: Clarendon Press.
- Hume, David. (2004/1748). *An enquiry concerning human understanding*. Mineola, N.Y.: Dover. First pub. 1748.
- Husserl, Edmund (1913). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*. Halle: Max Niemeyer.

- Husserl, Edmund (1913-1921/1900-1901). *Logische Untersuchungen*. 2 Bde. Haag: Niemeyer. First pub. 1900-1901.
- Husserl, Edmund (1928). Martin Heidegger (ed.). *Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins*. Halle: Max Niemeyer.
- Husserl, Edmund (1929). *Formale und transzendente Logik: Versuch einer Kritik der logischen Vernunft*. Halle-Salle: Max Niemeyer.
- Husserl, Edmund (1962/1936). *L'Origine de la géométrie*. Jacques Derrida (trans. and intro.). Paris: Presses Universitaires de France. Orig. pub. 1936.
- Husserl, Edmund (1964/1905-1910). *Leçons pour une phénoménologie de la conscience intime du temps*. Henri Dussort (trans.). Paris: Presses Universitaires de France. [Based on lectures originally given 1905-1910.]
- Husserl, Edmund (1965). Philosophy as rigorous science. Quentin Lauer (trans.). In *Phenomenology and the crisis of philosophy*, pp. 71-147. New York: Harper and Row.
- Husserl, Edmund (1965/1929). *Logique formelle et logique transcendantale: Essai d'une critique de la raison logique*. Suzanne Bachelard (trans.). Paris: Presses Universitaires de France. Orig. pub. in German in 1929.
- Isaye, Gaston (1953). La finalité de l'intelligence et l'objection kantienne. *Revue Philosophique de Louvain*, 51:29, 42-100.
- Isaye, Gaston (1954). La justification critique par rétorsion. *Revue Philosophique de Louvain*, 52, 205-233.
- Isaye, Gaston (1987). *L'Affirmation de l'être et les sciences positives*. Commentaire par Marc Leclerc, Préface par Jean Ladrière. Paris: Éditions Lethielleux.
- Jacobi, Friedrich Heinrich (1787). *David Hume über den Glauben, oder, Idealismus und Realismus*. Breslau, Prussia [now Poland]: G. Loewe.
- Jacobson, Arthur (1970). Russell and Strawson on referring. In Klemke (1970), pp. 285-308.
- James, William (1890). *The principles of psychology*. 2 vols. New York: H. Holt.
- James, William (1905/1897). *The sentiment of rationality*. New York: Longmans, Green & Co. First pub. 1897.
- James, William (1958/1904). A world of pure experience. In *Essays in radical empiricism and A pluralistic universe*, pp. 39-91. New York: Longmans, Green & Co. First pub. 1904 in *J. of Philosophy, Psychology, and Scientific Methods*, 1, 533-543, 561-570. Available online.
- James, William (1979/1897). *The will to believe and other essays in popular philosophy*. In Burkhardt, F. H. ; Bowers, F. ; & Skrupskelis, I. K. (eds.), *The*

*works of William James*, vol. 6, pp. 1-248. Cambridge, MA: Harvard Univ. Press.

Jammer, Max (1957). *Concepts of force: A study in the foundations of dynamics*. Cambridge, Mass.: Harvard Univ. Press.

Johnson Edward (1842). *Nuces philosophicae, or, the philosophy of things, as developed from the study of the philosophy of words*. London: Simpkin, Marshall & Co.

Johnstone, Henry W., Jr. (1964). Self-refutation and validity. *Monist*, 48:4, 467-485.

Johnstone, Henry W., Jr. (1970). *Philosophy and argument*. Penn. State Univ. Press.

Joseph, A. B. (1986). Focal central nervous system abnormalities in patients with misidentification syndromes. *Bibliotheca Psychiatrica*, 164, 68-79.

Kant, Immanuel & Beck, Lewis White. (1950/1783). *Prolegomena to any future metaphysics*. Indianapolis: Bobbs-Merrill. *Prolegomena* first pub. 1783.

Kant, Immanuel (1798-1799). *Essays and treatises on moral, political, and various philosophical subjects*. 2 vols. [English translator not given.] London: "printed for the translator and sold by William Richardson."

Kant, Immanuel (1900/1887). *Critique of pure reason*. J. M. D. Meiklejohn (trans.). London: George Bell & Sons. English trans. first pub. 1887.

Kant, Immanuel (1929/1770). *Inaugural dissertation and early writings on space*. John Handyside (trans.). London: Open Court. First pub. 1770.

Kant, Immanuel (1965/1929). *Critique of pure reason*. Normal Kemp Smith (trans.). New York: St. Martin's Press. English trans. first pub. 1929.

Kant, Immanuel (1974/1797). *Anthropology from a pragmatic point of view*. M. J. Gregor (trans.). Hague: Nijhoff. First pub. 1797.

Kant, Immanuel (1974/1800). *Logic*. Robert S. Hartman & Wolfgang Schwarz (trans.). New York: Bobbs-Merrill. [*Jäsche Logik* first pub. 1800.]

Kant, Immanuel (1997). *Correspondence*. Arnulf Zweig (trans. & ed.). Cambridge: Cambridge Univ. Press. [An updated and expanded edition of Kant, *Philosophical correspondence: 1755-95*, Arnulf Zweig (trans. & ed.), Univ. of Chicago Press, 1967 and 1970.]

Kapp, Reginald O. (1960). *Towards a unified cosmology*. London: Scientific Book Guild.

Kaufmann, Walter (1963). *The faith of a heretic*. Garden City, New York: Doubleday.

- Kempson, Ruth M. (1975). *Presupposition and the delimitation of semantics*. Cambridge: Cambridge Univ. Press.
- Kleene, Stephen Cole (1952). *Introduction to metamathematics*. New York: van Nostrand.
- Kleiner, Israel (2007). *A history of abstract algebra*. Birkhäuser Boston: Boston, MA.
- Klemke, E. D. (ed.) (1970). *Essays on Bertrand Russell*. Urbana, IL: Univ. of Illinois Press.
- Kline, George L. (1965). N. A. Vasil'ev and the development of many-valued logics. In Tymieniecka, Anna-Teresa (ed.), *Contributions to logic and methodology in honor of J. M. Bocheński*, pp. 315-325. Amsterdam: North-Holland.
- Kneale, William & Kneale, Martha (1962). *The development of logic*. Oxford: Clarendon.
- Kneser, Adolf (1918). *Mathematik und Natur von der Schwere: Zwei akademische Reden*. Breslau: Verlag von Trewendt & Granier.
- Knight, Thomas S. (1967). Questions and universals. *Philosophy and Phenomenological Research*, 27:4, 564-576.
- Koestler, Arthur (1967). *The ghost in the machine*. New York: Macmillan.
- Kokoszńska, Maria (1951). A refutation of the relativism of truth. *Studia Philosophica*, 4, 93-149.
- Koks, Don (2006). *Explorations in mathematical physics: The concepts behind an elegant language*. New York: Springer.
- Kordig, Carl R. (1970). Objectivity, scientific change, and self-reference. *Boston Studies in the Philosophy of Science*, 8, 519-523.
- Körner, Stephan (1970). *Categorical frameworks*. New York: Barnes & Noble.
- Kreyszig, E. (1991). *Differential geometry*. New York: Dover.
- Kripke, S. (1972). *Naming and necessity*. Cambridge, MA: Harvard Univ. Press.
- Kripke, Saul A. (1963). Semantical analysis of modal logic I. Normal modal propositional calculi. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9, 67-96.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, 28, 16-26.

- Kuhn, D. (2000). Theory of mind, metacognition, and reasoning: A life-span perspective. In Mitchell, P. & Riggs, K. J. (eds.), *Children's reasoning and the mind*, pp. 301-326. Hove, UK: Psychology Press.
- Kuhn, Helmut (1940). The phenomenological concept of "horizon." In Farber, Marvin (ed.), *Philosophical essays in memory of Edmund Husserl*, pp. 106-123. Cambridge, MA: Harvard Univ. Press.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480-498.
- Laczos, Cornelius (1974). *The Einstein decade, 1905-1915*. New York: Academic Press.
- Ladrière, Jean (1992/1957). *Les Limitations internes des formalismes: Étude sur la signification du théorème de Gödel et des théorèmes apparentes dans la théorie des fondements des mathématiques*. Sceaux, France: J. Gabay. First pub. 1957.
- Lambert, Johann Heinrich (1764). *Neues Organon oder Gedanken über die Erforschung und Bezeichnung des Wahren und dessen Unterscheidung vom Irrthum und Schein*. Leipzig: Wendler. 2 vols. Re-published in 1990, 2 vols., Berlin: Akademie-Verlag.
- Lambert, Karel (1969). Logical truth and microphysics. In Lambert, Karel (ed.), *The logical way of doing things*, pp. 93-117. New Haven, CT: Yale Univ. Press.
- Langendoen, D. Terence & Savin, Harris (1971). The projection problem for presuppositions. In Fillmore, C. & Langendoen, D. T. (eds.), *Studies in Linguistic Semantics*, pp. 373-388. New York: Holt, Reinhardt & Winston.
- Larson, R., & Segal, G. (1995). *Knowledge of meaning*. Cambridge, MA: MIT Press.
- Laszlo, Ervin (1972). *Introduction to systems philosophy: Toward a new paradigm of contemporary thought*. New York: Gordon and Breach.
- Lawrence, T.[Thomas] E.[Edward] (1935). *Seven pillars of wisdom: A triumph*. Garden City, NY: Doubleday, Doran & Co.
- Leibniz, Gottfried Wilhelm Freiherr von (1957). André Robinet (ed.). *Correspondance Leibniz-Clarke*. Paris: Presses Universitaires de France.
- Leibniz, Gottfried Wilhelm Freiherr von (1991/1693-1715). *De l'horizon de la doctrine humaine (1693); Apokatastasis panton (La restitution universelle) (1715). Textes inédits*. Michel Fichant (trans.). Paris: J. Vrin. *De l'horizon* dates from 1693; *La restitution universelle* dates from 1715.



- Lemmon, E. J. (1959). Is there only one correct system of modal logic? Part I. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 33, 23-40.
- Lemmon, E. J. (1971/1965). *Beginning logic*. London: Thomas Nelson. First pub. 1965.
- Lemmon, Edward J. (1956). Review of G.E.M. Anscombe's "Aristotle and the Sea Battle." *JSL*, 21, 388-389.
- Leopardi, Giacomo (1981). *A Leopardi reader*. Ottavio M. Casale (ed. & trans.). Urbana, IL: Univ. of Chicago Press.
- Lewin, Kurt (1920). *Der Ordnungstypus der genetischen Reihen in Physik, organismischer Biologie und Entwicklungsgeschichte*. Berlin: Borntrager.
- Lewin, Kurt (1920a). *Die Verwandtschaftsbegriffe in Biologie und Physik und die Darstellung vollständiger Stammbäume*. Berlin: Borntrager.
- Lewis, C. I. (1934). Experience and meaning. *Philosophical Review*, 43:2, 125-146.
- Lewis, Clarence Irving & Langford, Cooper Harold (1932). *Symbolic logic*. New York & London: Century Co.
- Lewis, David (1973). Causation. *J. of Philosophy*, 70:17, 556-567.
- Lewis, David. (1986). *On the plurality of worlds*. London & New York: Basil Blackwell.
- Lewis, Gilbert N. (1926). Light waves and light corpuscles. *Nature*, 117, 236-238.
- Lewis, Gilbert N. (1930). The symmetry of time in physics. *Science*, 71, 569-577.
- Libet, Benjamin (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8:4, 529-566.
- Libet, Benjamin; Gleason, Curtis A.; Wright, Elwood W.; & Pearl, Dennis K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain*, 106:3, 623-642.
- Lichtenberg, Georg Christoph (2012). *Georg Christoph Lichtenberg: Philosophical writings*. Steven Tester (ed. and trans.). Albany, NY: SUNY Press.
- Linsky, Leonard (1967). *Referring*. New York: Humanities Press.
- Locke, John (1690). *An essay concerning humane understanding*. London: Eliz. Holt for Thomas Basset.
- Lonergan, Bernard (1988/1967). *Collected works of Bernard Lonergan*. Frederick E. Crowe & Robert M. Doran (eds.). Vol. 4, 2nd edition. Toronto: Univ. of Toronto Press. First pub. 1967.
- Lorentz, Hendrik Anton (1913). *Das Relativitätsprinzip*. Leipzig: Teubner.

- Lorentz, Hendrik Anton. (1920/1919). *The Einstein theory of relativity: A concise statement*. New York: Brentano's. First pub. 1919.
- Lorenzen, Paul (1969). *Normative logic and ethics*. Mannheim/Zürich: Bibliographisches Institut.
- Lorenzen, Paul (1969a). *Einführung in die operative Logik und Mathematik*. Berlin: Springer.
- Loux, Michael J. (1979). Introduction: Modality and metaphysics. In Loux, Michael J. (ed.), *The possible and the actual: Readings in the metaphysics of modality*, pp. 15-63. Ithaca: Cornell Univ. Press.
- Lovejoy, A. O. (1922). The paradox of the thinking behaviorist. *Philosophical Review*, 31, 142-147.
- Lucas, J. R. (1970). *Freedom of the will*. Oxford: Clarendon Press.
- Łukasiewicz, Jan (1953). A system of modal logic. *J. of Computing Systems*, 1:3, 111-149.
- Łukasiewicz, Jan (1957/1951). *Aristotle's syllogistic from the standpoint of modern formal logic*. Oxford: Clarendon Press. 2nd ed.; first ed. 1951.
- Mach, Ernst (1900/1886). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen*. Jena: Fischer. First pub. 1886.
- MacIntyre, Alasdair (2010). On not knowing where you are going. *Proceedings and Addresses of the American Philosophical Association*, 84:2, 61-92.
- Mackie, J. L. (1962). The philosophy of John Anderson. *Australasian J. of Philosophy*, 40:3, 265-282.
- Mackie, J. L. (1964). Self-refutation: A formal analysis. *Philosophical Quarterly*, 14:56, 193-203. (Reprinted in Mackie, J. L., *Logic and knowledge*, pp. 54-67. Joan Mackie & Penelope Mackie (eds.) Oxford: Clarendon, 1985.)
- Mackie, J. L. (1973). *Truth, probability, and paradox: Studies in philosophical logic*. Oxford: Clarendon.
- Mackie, J. L. (1980). Truth and knowability. *Analysis*, 40, 90-92.
- Mandelbaum, Allen & Moser, Barry (1980). *The Divine Comedy of Dante Alighieri*. Berkeley, CA: Univ. of California Press.
- Marbach, Eduard (1993). *Mental representation and consciousness: Towards a phenomenological theory of representation and reference*. Dordrecht: Kluwer Academic Publishers.
- Maréchal, Joseph (1970). *A Maréchal reader*. Joseph Donceel (trans. and ed.). New York: Herder & Herder.

- Martin, John Levi (2011). *The explanation of social action*. Oxford: Oxford Univ. Press.
- Martin, Richard Milton (1953). On non-translational semantics. In *Actes du XIème Congres International de Philosophie*, Bruxelles 20-26 août 1953, vol. V: *Logic, Philosophical Analysis, Philosophy of Mathematics*, pp. 132-138. Amsterdam: North-Holland; Louvain: E. Nauwelaerts.
- Martin, Richard Milton (1958). *Truth and denotation: A study in semantical theory*. Chicago: Univ. of Chicago Press.
- Mason, Lucia & Scirica, Fabio (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction*, 16, 492-509.
- Maugham, W. Somerset (1950). *The summing up*. In Maugham, W. Somerset, *The Maugham reader*, pp. 177-682. Garden City, NY: Doubleday.
- Mauthner, Fritz (1901-1902). *Beiträge zu einer Kritik der Sprache* [*Contributions to a critique of language*]. 3 vols. Stuttgart: Cotta.
- McCall, Storrs (ed.) (1967). *Polish logic 1920-1939*. Oxford: Oxford Univ. Press.
- Meinong, Alexius (1960/1904). The theory of objects. Isaac Levi; D. B. Terrell; & R. M. Chisholm (trans.). In Chisholm, Roderick M. (ed.), *Realism and the background of phenomenology*, pp. 76-117. Glencoe, IL: Free Press. (Originally pub. 1904 as *Untersuchungen zur Gegenstandstheorie und Psychologie*. Leipzig: J.A. Barth.)
- Mellor, W. W. (1954). Believing the meaningless. *Analysis*, 15:2, 41-43.
- Menger, K. (1930). Der Intuitionismus. *Blätter deutscher Philosophie*, 4, 311-325.
- Mermin, N. David (2012). Commentary: Quantum mechanics: Fixing the shifty split. *Physics Today*, 65:7, 8-10.
- Meyerson, Émile (1925). *La déduction relativiste*. Paris: Payot.
- Meyerson, Émile (1985/1925). *The relativistic deduction: Epistemological implications of the theory of relativity*. David A. & Mary-Alice Sipfle (trans.). Boston: D. Reidel.
- Mill, J. S. (1971/1863). Of what sort of proof the principle of utility is susceptible. In Mill, J. S., *Utilitarianism*, Chap. 4. Indianapolis, IN: Bobbs-Merrill. Orig. pub. 1863.
- Misner, Charles W.; Thorne, Kip S.; and Wheeler, John Archibald. (1973). *Gravitation*. San Francisco, CA: W. H. Freeman.

- Moh, Shaw-kwei (1954). Logical paradoxes for many-valued systems. *JSL*, 19, 37-40.
- Moleski, Martin X. (1977). Retortion: The method of Gaston Isaye. *International Philosophical Quarterly*, 17:1, 59-83.
- Moore, G. E. (1965). A defense of common sense. In Ammerman, R. R. (ed.), *Classics of analytic philosophy*, pp. 47-67. New York: McGraw-Hill.
- Morgan, C. L. (1892/1885). *Springs of conduct*. London: Edward Arnold. First pub. 1885.
- Morris, Charles (1946). *Signs, language, and behavior*. New York: Prentice-Hall.
- Morrison, Roy D. (1994). *Science, theology and the transcendental horizon: Einstein, Kant and Tillich*. Atlanta, GA: Scholars Press.
- Mostowski, Andrzej (1950). Review of Helen Rasiowa's "A dziedziny logiki matematycznej. II. Logiki wielwartosciowe Łukasiewicza." [From the domain of mathematical logic. II. The many-valued logic of Łukasiewicz], *JSL*, 15, 223. [Rasiowa's original paper appeared in *Matematyka*, 3 (1950), 4-11.]
- Muck, Otto (1968). *The transcendental method*. Wm. D. Seidensticker (trans.). New York: Herder & Herder.
- Munro, Geoffrey (2011). [From an interview:] "Quick answers: Studying people who disbelieve studies." *Miller-McCune*, March/April, 13.
- Murawski, R. (1999). *Recursive functions and metamathematics: Problems of completeness and decidability, Gödel's theorems*. Dordrecht: Kluwer.
- Murray, Charles (2003). *Human accomplishment: The pursuit of excellence in the arts and sciences, 800 B.C. to 1950*. New York: HarperCollins.
- Myhill, J. R. (1950). A system which can define its own truth. *Fundamenta Mathematicae*, 37, 190-192.
- Natorp, Paul (1923/1910). *Die logischen Grundlagen der exakten Wissenschaften*. Leipzig and Berlin: Teubner. First pub. 1910.
- Nelson, Leonard (1970). *Progress and regress in philosophy from Hume and Kant to Hegel and Fries*. Vol. I. Julius Kraft (ed.), Humphrey Palmer (trans.). Oxford: Basil Blackwell.
- Nerlich, G. (1965). Presupposition and entailment. *American Philosophical Quarterly*, 2:1, 33-42.
- Newell, A.; Shaw, J.; & Simon, H. (1957). Empirical explorations of the logic theory machine: A case study in heuristic. *Proceedings Western Joint Computer Conf.*, XV, 218-230.

- Newton, Isaac (1934/1606), *Mathematical principles of natural philosophy and the system of the world*. Andrew Motte (trans.) (1729), rev. Florian Cajori. Berkeley: Univ. of California Press. Orig. printing 1606.
- Newton, Isaac (1952/1704). *Opticks, or a treatise on the reflections, refractions, inflections and colours of light*. New York: Dover. First pub. 1704.
- Nickerson, Raymond S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2:2, 175-220.
- Nietzsche, Friedrich (1909/1901). *The will to power: An attempted transvaluation of all values*. 2 vols. Anthony M. Ludovici & Oscar Levy (trans.). Edinburgh: T.N. Foulis. First German pub. 1901.
- Nietzsche, Friedrich (1918/1895). *The antichrist*. H. L. Mencken (trans.). New York: Alfred A. Knopf. First German pub. 1895.
- Nozick, Robert (2001). *Invariances: The structure of the objective world*. Cambridge, MA: Harvard Univ. Press.
- O'Leary, M. (2010). *Revolutions of geometry*. London: Wiley.
- Odegard, D. (1966). On weakening excluded middle. *Dialogue*, 5, 232-236.
- Ogden, C. K. & Richards, I. A. (1923). *The meaning of meaning: A study of the influence of language upon thought and of the science of symbolism*. New York: Harcourt, Brace & World.
- Palmer, Humphrey (1981). Do circular arguments beg the question? *Philosophy*, 56:217, 387-394.
- Palmer, Humphrey (1983). The transcendental fallacy. *Kant-Studien*, 74:4, 387-404.
- Palmer, Humphrey (1985). *Presupposition and transcendental inference*. New York: St. Martin's Press.
- Palmer, Humphrey (1994). Foundations, circularity, and transcendental arguments. In De brock, G. and Hulswit, M. (eds.), *Living Doubt: Essays concerning the epistemology of Charles Sanders Peirce*, pp. 33-41. Dordrecht, Holland: Kluwer Academic Publishers.
- Passmore, John (1961). *Philosophical reasoning*. London: Gerald Duckworth.
- Pauli, Wolfgang (ed.) (1955). *Niels Bohr and the development of physics*. London: Pergamon.
- Pearl, Judea (2009/2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge Univ. Press. First pub. 2000.

- Peirce, Charles Sanders (1931-1935). *Collected Papers of Charles S. Peirce*. Charles Hartshorne & Paul Weiss (eds.). Cambridge, MA: Harvard Univ. Press.
- Penrose, Roger (1999/1989). *The emperor's new mind: Concerning computers, minds and the laws of physics*. Oxford: Oxford Univ. Press.
- Perelman, Chaim & Tyteca, L. Olbrechts (1971). *New rhetoric: A treatise on argumentation*. Notre Dame, IN: Univ. of Notre Dame Press.
- Peters, W. S. (1968). Kants Verhältnis zu J. H. Lambert. *Kant-Studien*, 59, 448–453.
- Philipse, Herman (2009). Can philosophy be a rigorous science? *Royal Institute of Philosophy Supplement*, 65, 155-170.
- Pitchford, N. J. & Biggam, C. P. (eds.) (2006). *Progress in colour studies*. 2 vols. Philadelphia: John Benjamins Publishing.
- Planck, Max (1933). *Where is science going?* James Murphy (trans.). London: G. Allen & Unwin.
- Plantinga, A. (1974). *The nature of necessity*. Oxford: Oxford Univ. Press.
- Plantinga, A. (1976). Actualism and possible worlds. *Theoria*, 42, 139–60.
- Plato (1952). *Theaetetus*. Benjamin Jowett (trans.). In *The dialogues of Plato, Great books of the western world*, vol. 7, pp. 512-550. Chicago, IL: Encyclopaedia Britannica.
- Polányi, Michael (1952). The stability of beliefs. *British J. for the Philosophy of Science*, 3:11, 217-232.
- Polányi, Michael (1959). *The study of man*. Chicago: Univ. of Chicago Press.
- Pólya, George (1941). Heuristic reasoning and the theory of probability. *American Mathematical Monthly*, 48, 450-465.
- Pólya, George (1950). Let us teach guessing. In Gonseth, Ferdinand, *Études de Philosophie des Sciences, en hommage à Ferdinand Gonseth à l'occasion de son soixantième anniversaire*, pp. 147-154. Neuchâtel: Éditions du Griffon.
- Pólya, George (1962-65). *Mathematical discovery: On understanding, learning and teaching problem solving*. Vol. I (1962); Vol. II (1965). New York: John Wiley.
- Pólya, George (1968/1954). *Mathematics and plausible reasoning*. Vol. I: *Induction and analogy in mathematics*; Vol. II, *Patterns of plausible inference*. Princeton, NJ: Princeton Univ. Press. First pub. 1954.
- Pólya, George (1971/1945). *How to solve it: A new aspect of mathematical method*. Princeton, NJ: Princeton Univ. Press. First pub. 1945.

- Popper, Karl (1963). *Conjectures and refutations: The growth of scientific knowledge*. New York: Harper & Row.
- Presley, C. F. (1961). Arguments about meaninglessness. *British J. for the Philosophy of Science*, XII:47, 229-230.
- Preyer, G. & Peter, G. (eds.) (2005). *Contextualism in philosophy*, Oxford: Oxford Univ. Press.
- Priest, Graham (1987). *In contradiction: A study of the transconsistent*. Dordrecht: Martinus Nijhoff. (Second edition, Oxford: Oxford Univ. Press, 2006.)
- Priest, Graham (1991). The limits of thought—and beyond. *Mind*, New Series, 100:3, 361-370.
- Priest, Graham (1995/2002). *Beyond the limits of thought*. Cambridge: Cambridge Univ. Press. Second expanded edition, Oxford: Oxford Univ. Press, 2002.
- Priest, Graham (1996). Some priorities of Berkeley. In Copeland, B. J. (ed.), *Logic and reality: Essays on the legacy of Arthur Prior*, pp. 480-487. Oxford: Clarendon.
- Priest, Graham (2006). *In contradiction: A study of the transconsistent*. Oxford: Clarendon.
- Priest, Graham (2007). Paraconsistency and dialetheism. In Gabbay, Dov M. & Woods, John (eds.), *Handbook of the history of logic: Volume 8: The many valued and nonmonotonic turn in logic*, pp. 129-204. Amsterdam: Elsevier.
- Priest, Graham (2016). *Towards non-being: The logic and metaphysics of intentionality*. Oxford: Oxford Univ. Press.
- Priest, Graham; Beall, JC; & Armour-Garb, B. (eds.) (2004). *The law of non-contradiction*. Oxford: Oxford Univ. Press.
- Priest, Graham; Routley, R.; & Norman, J. (eds.) (1989). *Paraconsistent logic: Essays on the inconsistent*. München: Philosophia Verlag.
- Prior, Arthur N. (1952). In what sense is modal logic many-valued? *Analysis*, 12, 138-143.
- Prior, Arthur N. (1953). Three-valued logic and future contingents. *Philosophical Quarterly* (St. Andrews), 3, 317-326.
- Prior, Arthur N. (1955). Many-valued and modal systems: An intuitive approach. *Philosophical Review*, 64, 626-630.
- Prior, Arthur N. (1955a). *Formal logic*. Oxford: Oxford Univ. Press.
- Prior, Arthur N. (1957). *Time and modality*. Oxford: Clarendon.

- Prior, Arthur N. (1957a). The necessary and the possible: The first of three talks on "The Logic Game." *The Listener*, 57, 627-628. Symbolism and analogy: The second of three talks on "The Logic Game." *The Listener*, 57, 675-678. Many-valued logics: The last of three talks on "The Logic Game." *The Listener*, 57, 717-719.
- Prior, Arthur N. (1959). Thank goodness that's over. *Philosophy*, 34, 12-17.
- Prior, Arthur N. (1961). On a family of paradoxes. *NDJFL*, 2, 16-32.
- Prior, Arthur N. (1962). The formalities of omniscience. *Philosophy*, 37, 114-129.
- Prior, Arthur N. (1966). Two essays on temporal realism. In Copeland, B. J. (ed.), *Logic and reality: Essays on the legacy of Arthur Prior*, pp. 43-51. Oxford: Clarendon.
- Prior, Arthur N. (1967). *Past, present, and future*. Oxford: Clarendon.
- Prior, Arthur N. (1968). *Papers on time and tense*. Oxford: Clarendon.
- Prior, Arthur N. (1968a). Egocentric logic. *Nous*, 2, 191-207.
- Putnam, Hilary (1975). The meaning of 'meaning'. In Gunderson, Keith (ed.), *Language, mind, and knowledge*, pp. 131-193. Minnesota Studies in the Philosophy of Science, vol. 7. Minneapolis: Univ. of Minnesota Press.
- Pylyshyn, Z. (1984). *Computation and cognition: Towards a foundation for cognitive science*. Cambridge, MA: MIT Press.
- Quine, W. V. O. (1948). On what there is. *Rev. of Metaphysics*, 2:3, 21-38.
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MA.: Technology Press.
- Quine, W. V. O. (1963/1953). *From a logical point of view*. New York: Harper and Row. Orig. pub. 1953.
- Quine, W. V. O. (1969). *Ontological relativity and other essays*. New York: Columbia Univ. Press.
- Quine, W. V. O. (1978/1977). Facts of the matter. In Shahan, Robert W. & Swoyer, Chris (eds.), *Essays on the philosophy of W. V. Quine*, pp. 155-169. Norman, OK: Univ. of Oklahoma Press. First pub. 1977.
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
- Recorde, Robert (2012/1556). *The castle of knowledge: A facsimile of the first edition imprinted at London by Reginalde Wolfe, 1556*. Mickleover, Derby, U.K.: Renascent Books. First pub. 1556.
- Rehg, William R. (1989). Lonergan's performative transcendental argument against scepticism. *Proceedings of the American Catholic Philosophical Association*, 63, 257-268.



- Reichenbach, Hans (1944). *Philosophic foundations of quantum mechanics*. Berkeley, CA: Univ. of California Press.
- Reichenbach, Hans (1947). *Elements of symbolic logic*. New York: Macmillan.
- Reichenbach, Hans (1958/1928). *The philosophy of space and time*. Maria Reichenbach (trans.). New York: Dover. (Orig. pub. 1928 as *Philosophie der Raum-Zeit-Lehre*. Berlin: Gruyter.)
- Reichenbach, Hans (1965/1920). *The theory of relativity and a priori knowledge*. Maria Reichenbach (trans. and ed.). Berkeley, CA: Univ. of California Press. (Orig. pub. 1920 as *Relativitätstheorie und Erkenntnis Apriori*. Berlin: Springer.)
- Rescher, Nicholas & Urquhart, Alasdair (1971). *Temporal logic*. Wien/New York: Springer-Verlag.
- Rescher, Nicholas (1962). Quasi-truth-functional systems of propositional logic. *JSL*, 27, 1-10.
- Rescher, Nicholas (1963). An interpretation of Aristotle's doctrine of future contingency and excluded middle. In Rescher, Nicholas, *Studies in the history of Arabic logic*, pp. 43-54. Pittsburgh, PA: Univ. of Pittsburgh Press.
- Rescher, Nicholas (1964). A note on self-referential statements. *NDJFL*, 5, 218-220.
- Rescher, Nicholas (1965). An intuitive interpretation of systems of four-valued logic. *NDJFL*, 6, 154-156.
- Rescher, Nicholas (1969). *Many-valued logic*. New York: McGraw-Hill.
- Rescher, Nicholas (1975). *A theory of possibility: A constructivistic and conceptualistic account of possible individuals and possible worlds*. Pittsburgh: Univ. of Pittsburgh Press.
- Rescher, Nicholas (1979). Philosophical disagreement: An essay towards orientational pluralism in metaphilosophy. *Rev. of Metaphysics*, 32:2, 217-251.
- Rescher, Nicholas (2005). *Epistemic logic: A survey of the logic of knowledge*. Pittsburgh: Univ. of Pittsburgh Press.
- Rescher, Nicholas (2009). *Unknowability: An inquiry into the limits of knowledge*. New York: Lexington.
- Rescher, Nicolas (2014). Meaninglessness. In Rescher, Nicolas, *Logical inquiries: Basic issues in philosophical logic*, pp. 63-66. Berlin, Germany: De Gruyter.
- Richards, Robert J. (1989). *Darwin and the emergence of evolutionary theories of mind and behavior*. Chicago: Univ. of Chicago Press.

- Richman, R. J. (1953). On the self-reference of a meaning theory. *Philosophical Studies*, 4, 69-72.
- Roberts, G. (1991). Delusional belief systems and meaning in life: A preferred reality. *Brit. J. of Psychiatry*, 161, 298-308.
- Robinson, Richard (1962/1950). *Definition*. Oxford: Oxford Univ. Press. First pub. 1950.
- Romanes, George (1884). Mr. Lloyd Morgan on instinct. *Nature*, Feb. 21, 379-381.
- Romanes, George John (1883). *Animal intelligence*. New York: Appleton.
- Romanes, George John (1888). *Mental evolution in man: Origin of human faculty*. New York: Appleton.
- Romanes, George John (1895/1883). *Mental evolution in animals*. New York: Appleton. First pub. 1883.
- Rorty, Richard (1961). The limits of reductionism. In Lieb, J. (ed.), *Experience, existence, and the good*, pp. 100-116. Carbondale: Southern Illinois Univ. Press.
- Rosenthal-Schneider, I. (1980). *Reality and scientific truth*. Detroit: Wayne State Univ. Press.
- Routley, R & Routley, V. (1975). The role of inconsistent and incomplete theories in the logic of belief. *Communication & Cognition*, 8, 185-235.
- Routley, Richard & Routley, Valerie (1973). Rehabilitating Meinong's theory of objects. *Revue Internationale de Philosophie*, 27:104/105 (2/3), 224-254.
- Routley, Richard (1980). *Exploring Meinong's jungle and beyond: An investigation of noneism and the theory of items*. Departmental monograph number three, Philosophy Department, Research School of Social Sciences, Australian National University, Canberra, Australia. A complete PDF copy of the original book can be downloaded from Digitalcommons. [NB: The PDF file is approximately one-half GB.]
- Routley, Richard (2010/1981). Necessary limits to knowledge: Unknowable truths. *Synthese*, 173:1, 107-122. First pub. 1981.
- Royce, Josiah (1904). The eternal and the practical. *Philosophical Review*, 13, 128-129.
- Royce, Josiah (1919). *Lectures on modern idealism*. New Haven, CT: Yale Univ. Press.
- Royce, Josiah (1959/1899). *The world and the individual*. New York: Dover. First pub. 1899.

- Ruesch, Jurgen & Bateson, Gregory (1968). *Communication: The social matrix of psychiatry*. New York: W. W. Norton.
- Russell, Bertrand (1903). *The principles of mathematics*. Cambridge Univ. Press.
- Russell, Bertrand (1912-13). On the notion of cause. *Proceedings of the Aristotelian Society, XIII*, 1-26.
- Russell, Bertrand (1921). *The analysis of mind*. London: Allen & Unwin.
- Russell, Bertrand (1940). *An inquiry into meaning and truth*. London: G. Allen & Unwin.
- Russell, Bertrand (1950/1903). *The principles of mathematics*. London: George Allen & Unwin.
- Russell, Bertrand (1959). "Face to Face," interview with John Freeman, BBC interviews. Available online.
- Russell, Bertrand (1961/1922). Introduction. In Wittgenstein (1961/1921), pp. ix-xxii). First pub. 1922.
- Russell, Bertrand (1967/1919). *Introduction to mathematical philosophy*. London: Allen & Unwin. First pub. 1919.
- Russell, Bertrand (1972/1914). *Our knowledge of the external world as a field for scientific method in philosophy*. London: George Allen & Unwin. First pub. 1914.
- Rüstow, Alexander (1980). *Freedom and domination: A historical critique of civilization*. S. Attanasio (trans.). Princeton, N.J.: Princeton Univ. Press.
- Ryle, Gilbert (1959/1946). Philosophical arguments. In Ayer, A. J. (ed.), *Logical positivism*, pp. 327-344. New York: Free Press. First pub. 1946.
- Saari, Heikki (1991). Some aspects of R. G. Collingwood's doctrine of absolute presuppositions. *International Studies in Philosophy, XXIII*:1, 61-73.
- Salerno, J., (ed.) (2009). *New essays on the knowability paradox*. Oxford: Oxford Univ. Press.
- Salovey, Peter & Mayer, John D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9*:3, pp. 185-211.
- Schiffer, S. (1972). *Meaning*. Oxford: Oxford Univ. Press.
- Schiffer, Stephen (1978). The basis of reference. *Erkenntnis, 13*:1, 171-206.
- Schilpp, Paul Arthur (ed.) (1959/1949). *Albert Einstein: Philosopher-scientist*. 2 vols. New York: Harper & Row. First Pub. 1949.
- Schlenker, Philippe (2007). Anti-dynamics: Presupposition projection without dynamic semantics. *J. of Logic, Language and Information, 16*:3, 325-256.

- Schlenker, Philippe (2008). Be articulate: A pragmatic theory of presupposition. *Theoretical Linguistics*, 34, 157–212.
- Schlenker, Philippe (2009). Local contexts. *Semantics and Pragmatics*, 2:3: 1–78.
- Schlesinger, George N. (1985). *The range of epistemic logic*. Atlantic Highlands, NJ: Aberdeen Univ. Press.
- Schlick, Moritz (1925/18). *Allgemeine Erkenntnislehre*. Berlin: Julius Springer. First pub. 1918.
- Schlick, Moritz (1936). Meaning and verification. *Philosophical Review*, XLV:4, 339-369.
- Schlick, Moritz (1949/1931). Causality in everyday life and in recent science. In Feigl, Herbert & Sellars, Wilfrid (eds.), *Reading in philosophical analysis*, pp. 515-533. New York: Appleton-Century-Crofts. First German pub. 1931.
- Schlick, Moritz (1959/1932-33). Positivism and realism. David Rynin (trans). In Ayer, A. J. (ed.), *Logical positivism*, pp. 82-107. New York: Free Press. First pub. 1932-33.
- Schmidt, Paul F. (1957). Self-referential justification. *Philosophical Studies*, 8, 49-54.
- Schopenhauer, Arthur (1909/1818-1819). *The world as will and idea*. R. B. Haldane & J. Kemp (trans.). 3 vols. London: Kegan Paul, Trench, Trübner & Co.
- Schulze, Gottlob Ernst (1792). *Aenesidemus: oder, Über die Fundamente der von dem Herrn Professor Reinhold in Jena gelieferten Elementar-Philosophie*. Helmstädt, Germany: [anonymous publisher].
- Schweitzer, Albert (1949/1933). *Out of my life and thought: An autobiography*. C. T. Campion (trans.). New York: Henry Holt.
- Sellars, W. (1954). Presupposing. *Philosophical Review*, 63, 197–215.
- Sesonske, Alexander & Fleming, Noel (eds.) (1965). *Meta-meditations: Studies in Descartes*. Los Angeles, CA: Wadsworth.
- Sesonske, Alexander (1966). Plato's apology: *Republic I*. In Sesonske, Alexander (ed.), *Plato's Republic: Interpretation and criticism*, pp. 40-47. Los Angeles, CA: Wadsworth.
- Shaw, George Bernard (1934). *Too true to be good, village wooing & on the rocks. Three plays*. London: Constable.
- Simmonds, J. G. (1994/1982). *A brief on tensor analysis*. New York: Springer-Verlag. 2nd ed. First pub. 1982.

- Simons, Mandy (2001). On the conversational basis of some presuppositions. In Hastings, R.; Jackson, B.; & Zvolensky, Z. (eds.), *Proceedings of Semantics and Linguistics Theory 11*, pp. 431-448. Ithaca, NY: CLC Publications.
- Simons, Mandy (2003). Presupposition and accommodation: Understanding the Stalnakerian picture. *Philosophical Studies*, 112:3, 251-278.
- Simons, Mandy (2004). Presupposition and relevance. In Szabó, Zoltán (ed.), *Semantics vs. Pragmatics*, pp. 329-355. Oxford: Oxford Univ. Press.
- Simons, Mandy (2006). Foundational issues in presupposition. *Philosophy Compass*, 1:4: 357-372.
- Simons, Mandy (2007). Observations on embedding verbs, evidentiality, and presupposition. *Lingua*, 117:6, 1034-1056.
- Sims, Andrew C.P. (2003). *Symptoms in the mind: An introduction to descriptive psychopathology*. Edinburgh: Saunders.
- Singer, Isaac Bashevis (1982). *The collected stories of Isaac Bashevis Singer*. New York: Farrar, Straus, Giroux.
- Skinner, B. F. (1966). An operant analysis of problem-solving. In Kleinmuntz, B. (ed.), *Problem solving – Research, method, and theory*, pp. 225-257. New York: Wiley.
- Smith, Barry (1995). Frege and Chomsky: Sense and psychologism. In Biro, John & Kotatko, Petr (eds.), *Frege: Sense and reference one hundred years later*, pp. 25-46. Dordrecht, Holland: Kluwer Academic Publishers.
- Smith, P. (2007). *An introduction to Gödel's theorems*. Cambridge: Cambridge Univ. Press.
- Smullyan, Raymond M. (1957). Languages in which self-reference is possible. *JSL*, 22, 55-67. [Reprinted in Jaakko Hintikka (ed.), *The philosophy of mathematics*, pp. 64-77. Oxford: Oxford Univ. Press.]
- Soames, S. (2003). *Philosophical analysis in the Twentieth Century*. Vol. 2: *The age of meaning*. Princeton, NJ: Princeton Univ. Press.
- Soames, S. (2010). *Philosophy of language*. Princeton, NJ: Princeton Univ. Press.
- Solomaa, Arto (1959). On many-valued systems of logic. *Ajatus*, 22, 115-159.
- Spaulding, E. G. (1918). *The new rationalism*. New York: Holt.
- Spearman, Charles E. (1927). *The abilities of man*. New York: Macmillan.
- Stalnaker, Robert (1972). Pragmatics. In Davidson, D. & Harman, G. (eds.), *Semantics of Natural Language*, pp. 389-408. Reidel: Dordrecht.
- Stalnaker, Robert (1973). Presuppositions. *J. of Philosophical Logic*, 2, 447-457.

- Stalnaker, Robert (1974). Pragmatic presuppositions. In Munitz, M. & Unger, P. (eds.), *Semantics and Philosophy*, pp. 197-214. New York: New York Univ. Press.
- Stalnaker, Robert (1998). On the representation of context. *J. of Logic, Language and Information*, 7, 3–19.
- Stanovich, K. E. & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *J. of Educational Psychology*, 89, 342-357.
- Stanovich, K. E. (1999). *Who is rational? Studies in individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *J. of Experimental Psychology: General*, 127, 161-188.
- Stapleford, Scott (2008). *Kant's transcendental arguments: Disciplining pure reason*. London: Continuum Internatl. Publishing Group.
- Stapp, Henry (1971). S-matrix interpretation of quantum theory. *Physical Review D*, 3:6, 1303-1320.
- Stebbing, Lizzie Susan (1933/1930). *A modern introduction to logic*. London: Methuen. First pub. 1930.
- Stephani, Hans and Stewart, John (1990). *General Relativity*. Cambridge: Cambridge Univ. Press.
- Stern, Robert (ed.) (1999). *Transcendental arguments: Problems and prospects*. Oxford: Oxford Univ. Press.
- Stern, Robert (ed.) (2000). *Transcendental arguments and scepticism: Answering the question of justification*. Oxford: Clarendon.
- Strang, Colin (1960). Aristotle and the sea battle. *Mind*, 69, 447-465.
- Strawson, P. F. (1952). *Introduction to logical theory*. London: Methuen.
- Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. London: Methuen.
- Strawson, P. F. (1970/1950). On referring. In Klemke (1970), pp. 146-172. [Reprinted from *Mind*, 59:235, 1950, 320-344.]
- Strawson, P. F. (1970a/1954). A reply to Mr. Sellars. In Klemke (1970), 190-204. [Reprinted from *Philosophical Review*, 63:2, 1954, 216-231.]
- Strawson, P. F. (1970b/1964). Identifying reference and truth-values. In Klemke (1970), pp. 236-255. [Reprinted from *Theoria*, 30, 1964, 96-118.]

- Stroll, Avrum (1955). Believing the meaningless: A reply to Mr. Mellor. *Analysis*, 16:2, 45-48.
- Stroud, Barry (1968). Transcendental arguments. *J. of Philosophy*, 65, 241- 256.
- Suber, Peter (1987). A bibliography of works on reflexivity. In Bartlett and Suber (1987), pp. 259-362.
- Suber, Peter (1994). Question-begging under a non-foundational model of argument. *Argumentation*, 8, 241-50.
- Sullivan, J. W. N. (1934). *Contemporary mind: Some modern answers*. London: H. Toulmin.
- Suppes, Patrick (1969/1950). Measurement, empirical meaningfulness, and three-valued logic. In Suppes, Patrick, *Studies in the methodology and foundations of science: Selected papers from 1951 to 1969*, pp. 65-79. Dordrecht-Holland: D. Reidel. [Reprinted from C. W. Churchman & P. Rattoosh (eds.), *Measurement: Definitions and theories*, pp. 129-143. New York: John Wiley, 1950.]
- Taub, A. H. (1961). Quantum logics (strict- and probability-logics). [A review of an unfinished manuscript which was written about 1937, by John von Neumann.] In von Neumann, John, *John von Neumann: Collected Works*, pp. 195-197. Oxford: Pergamon, 1962.
- Taylor, C. W. (1958). Some variables functioning in productivity and creativity. In Taylor, C. W. (ed.), *The second University of Utah research conference on the identification of scientific talent*, pp. 4-7. Salt Lake City: Univ. of Utah Press.
- Taylor, C. W. (ed.) (1964). *Creativity: Progress and potential*. New York: McGraw-Hill.
- Taylor, J. R. (1982). *An introduction to error analysis*. Oxford: Oxford Univ. Press.
- Taylor, Richard (1957). The problem of future contingencies. *Philosophical Review*, 66, 1-28.
- Thorndike, E. L. (1920). Intelligence and its uses. *Harper's Magazine*, 140, 227-235.
- Todd, William (1968). *Analytical solipsism*. Hague: Martinus Nijhoff.
- Tolstoy, Leo [Tolstoi, Lyof N.] (1899). *What is art?* Aylmer Maude (trans.). New York: Thomas Y. Crowell.
- Tolstoy, Leo [Tolstoi, Lyof N.] (1919). *Anna Karenina*. Constance Garnett (trans.). Philadelphia, PA: George W. Jacobs.

- Tooley, Michael (2011). The skeptical challenges of Hume and Berkeley: Can they be answered? *Proceedings and Addresses of the American Philosophical Association*, 85:2, 27-46.
- Toribio, Josefa & Clark, Andy (eds.) (1998). *Language and meaning in cognitive science: Cognitive issues and semantic theory, artificial intelligence and cognitive science: Conceptual issues, vol. 4*. New York: Garland.
- Toulmin, Stephen (1969). Ludwig Wittgenstein. *Encounter*, 32, January, 58-71.
- Urban, W. M. (1929). *The intelligible world*. London: Allen & Unwin.
- Urban, W. M. (1949). *Beyond realism and idealism*. London: Allen & Unwin.
- van Fraassen, Bas C. (1968). Presupposition, implication, and self-reference. *J. of Philosophy*, 65, 136-152.
- van Fraassen, Bas C. (1969). Presuppositions, supervaluations, and free logic. In Lambert, Karel (ed.), *The logical way of doing things*, pp. 67-91. New Haven, CT: Yale Univ. Press.
- van Fraassen, Bas C. (1971). *Formal semantics and logic*. New York: Macmillan.
- van Inwagen, Peter (2006). *The problem of evil*. Oxford: Clarendon Press.
- Vasil'ev, Nikolai A. (1924). Imaginary (non-Aristotelian) logic. *Atti del V Congresso Internazionale di Filosofia* (Napoli, 5-9 maggio). Napoli, Italy: Società Anonima Editrice Francesco Perrella.
- Volckelt, Johannes (1918). *Gewissheit und Wahrheit*. München: Beck.
- von Neumann, John (1955/1932). *Mathematical foundations of quantum mechanics*. Princeton, NJ: Princeton Univ. Press. [First pub. as *Mathematische Grundlagen der Quantenmechanik*. Berlin: Julius Springer, 1932.]
- von Schubert-Soldern, Richard (1884). *Grundlagen einer Erkenntnistheorie*. Leipzig: R. Reisland.
- von Weizsäcker, C. F. (1971). The unity of physics. In Bastin, Ted (ed.), *Quantum theory and beyond*, pp. 229-262. Cambridge: Cambridge Univ. Press.
- von Weizsäcker, C. F. (1973). Classical and quantum description. In Mehra, J. (ed.), *The physicist's conception of nature*, pp. 635-667. Dordrecht, Holland: D. Reidel.
- von Weizsäcker, Carl Friedrich (1980/1971). *The unity of nature*. Francis J. Zucker (trans.). New York: Farrar, Straus, Giroux. First pub. in German 1971.
- von Wright, G. H. (1952). Interpretations of modal logic. *Mind*, 61:242, 165-177.



- Walsh, Roger & Vaughan, Frances (1983). Towards an integrative psychology of well-being. In Walsh, Roger & Shapiro, Deane H. (eds.), *Beyond health and normality: Explorations of exceptional psychological well-being*, pp. 388-431. New York: Van Nostrand.
- Wang, Hao (1953). Review of F.B. Fitch, *Symbolic logic: An introduction*. *J. of Philosophy*, 50:4, 129-134.
- Wang, Hao (1987). *Reflections on Kurt Gödel*. Cambridge, MA.: MIT Press.
- Ward, Lester Frank (1907). *Pure sociology: A treatise on the origin and spontaneous development of society*. New York: Macmillan.
- Watanabe, S. (1985). *Pattern recognition: Human and mechanical*. New York: Wiley.
- Watanabe, Satosi (1965). Une explication mathématique du classement d'objets. In Dockx, S. & Bernays, P. (eds.), *Information and prediction in science: Proceedings of a symposium sponsored by the Académie Internationale de Philosophie des Sciences, Brussels 3-8, September, 1962*, pp. 39-76. New York: Academic Press.
- Watanabe, Satosi (1969). *Knowing and guessing: A quantitative study of inference and information*. New York: John Wiley.
- Watanabe, Satosi (1986). Epistemological relativity. *Annals of the Japan Association for Philosophy of Science*, 7:1, pp. 1-14.
- Watzlawick, Paul; Beavin, Janet Helmick; & Jackson, Don D. (1967). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. New York: W. W. Norton.
- Weiler, Gershon (1958). On Fritz Mauthner's critique of language. *Mind*, New Series, 67:265, 80-87.
- Weiler, Gershon (1970). *Mauthner's critique of language*. Cambridge Univ. Press.
- Weinstock, M., & Cronin, M. A. (2003). The everyday production of knowledge: Individual differences in epistemological understanding and juror-reasoning skill. *Applied Cognitive Psychology*, 17, 161-181.
- Weinstock, Michael (2009). Relative expertise in an everyday reasoning task: Epistemic understanding, problem representation, and reasoning competence. *Learning and Individual Differences*, 19:4, 423-434.
- Weinstock, Michael; Neuman, Yair; & Tabak, Iris (2004). Missing the point or missing the norms? Epistemological norms as predictors of students' ability to identify fallacious arguments. *Contemporary Educational Psychology*, 29:1, 77-94.

- Wheeler, John A. (1985). Physics in Copenhagen in 1934 and 1935. In French, A. P. & Kennedy, P. J. (eds.), *Niels Bohr: A centenary volume*, pp. 221-226. Cambridge, MA: Harvard Univ. Press.
- Wheeler, John A. (1990). Information, physics, quantum: The search for links. In Zurek, Wojciech Hubert (ed.), *Complexity, entropy, and the physics of information*, pp. 4-18. Redwood City, CA: Addison-Wesley.
- Whitehead, Alfred North (1961/1922). The principle of relativity with applications to physical science. In Northrup, F. S. C. & Gross, Mason W. (eds.), *Alfred North Whitehead: An anthology*, pp. 295-357. New York: Macmillan. First pub. 1922.
- Williamson, Timothy (2000). *Knowledge and its limits*. Oxford Univ. Press.
- Williamson, Timothy (2007). *The philosophy of philosophy*. Oxford: Blackwell.
- Wisdom, John (1953). *Philosophy and psycho-analysis*. Oxford: Blackwell.
- Wittgenstein, L. (1953). *Philosophical investigations*. G. E. M. Anscombe (trans.). New York: MacMillan.
- Wittgenstein, Ludwig (1956). *Remarks on the foundations of mathematics*. In von Wright, G. H.; Rhees, R.; & Anscombe, G. E. M. (eds.); Anscombe, G. E. M. (trans.). Oxford: Basil Blackwell.
- Wittgenstein, Ludwig (1961/1921). *Tractatus Logico-philosophicus*. Pears, D. F. & McGuinness, B. F. (eds.) New York: Humanities Press. (First pub. in German in *Annalen der Naturphilosophie*, 1921.)
- Wolf, Susan (2011). Good-for-nothings. *Proceedings and Addresses of the American Philosophical Association*, 85:2, 47-64.
- Woodfield, Andrew (ed.) (1982). *Thought and object: Essays on intentionality*. Oxford: Clarendon.
- Woods, J. (2003). *Paradox and paraconsistency*. Cambridge: Cambridge Univ. Press.
- Zaner, Richard M. (1968). Philosophy and rhetoric. *Philosophy & Rhetoric*, 1, 61-77.
- Zermelo, Ernst (1967/1908) The possibility of a well-ordering. In Heijenoort, Jean van (ed.), *From Frege to Gödel: A source book in mathematical logic, 1879-1931*, pp. 181-198. Cambridge, MA: Harvard Univ. Press. First pub. 1908.
- Zinov'ev, Aleksander Aleksandrovich (1963-64). Two-valued and many-valued logic. *Soviet Studies in Philosophy*, 2, 69-84.

# Index

- absolute
  - length, 259, 283
  - simultaneity, 259, 283
- absolute presuppositions
  - Collingwood, 130
- absolute space, 128, 449-451, 454-457, 538, 560, 564, 618, 659
- absolute time, 128, 446, 455, 457, 538, 540, 560, 564, 659
- absolute truth
  - as framework-independent, 129
- abstract objects, 60, 298, 370, 494
- abstraction
  - Collingwood's summit of the pyramid, 71
- accelerated reference frames. *See* non-inertial reference frames.
- accessibility relations
  - between possible worlds, 139
- action
  - forms of in physics, 516
  - in a continuous medium, 516
  - at-a-distance, 516
  - by-contact, 516
- acts
  - illocutionary, 92
  - locutionary, 92
  - perlocutionary, 92
- Akademie Olympia, 534
- Albrecht, Andreas, 542
- allegory of the cave (Plato), 287
- ambiguity, 209, 239
- American Philosophical Association, 6-7
- analysis
  - retrospective application of, 253
- Anderson, John, 113, 115, 479
  - unspeakability, 113
- Andreasen, Nancy, 648, 650, 654
- Angell, R. B., 411
- Anna Karenina* (Tolstoy), 668
- anosognosia*, 264
- anthropology, 193
- Anthropology from a Pragmatic Point of View* (Kant), 648
- Aquinas, St. Thomas, 120, 761
- Åqvist, Lennart, 249, 267, 276, 268
- Archimedes of Syracuse, 317
- argument. *See also* under argumentation.
  - ad hominem*, 105, 121, 126
  - critical vs. constructive, 110
  - mathematical
    - nature of, 107
  - metalogical
    - transcendental basis, 126
  - rhetorical, 126
  - self-referential, 127
- argumentation
  - by rétorsion. *See* rétorsion.
  - conceptual, 125
  - framework-relative, 107
  - metalogical, 126-129
  - performative. *See* argumentation, pragmatism. *See also* argumentation, self-referential, pragmatism.
  - pragmatical
    - uses of, 109
  - rhetorical, 123, 125
    - performative nature of, 106
    - distinct from conceptual, 125
  - self-referential
    - metalogical, 192
    - pragmatical, 110-111, 127, 192
  - transcendental, 130-131, 145. *See*

- also* transcendental arguments.
- Aristotle, 120  
 four causes, 459  
*Organon*, 633
- Armour-Garb, B., 275
- artificial intelligence, 691  
 symbolic, 691
- Aster, Ernst, 491
- Augustine of Hippo (St. Augustine), 51
- Austin, John, 92
- autarky, 338
- autoconcordance, 352  
 in Husserl, 336
- autonomous use, 13, 93
- Avenarius, Richard Heinrich Ludwig, 534
- axiomatic method, 342
- Ayer, A. J., 234, 266
- Azande people, 652
- Bach, Kent, 67
- Bachelard, Gaston, 670
- Bachelard, Suzanne, 728
- Bacon, Francis, 633
- baseless beliefs, 37, 40-43, 639, 758  
 primitive level of, 42  
 wincing mentally in reaction to, 41
- Beall, JC, 275
- behavioral economics, 355, 741
- behaviorism, 110
- beliefs  
 baseless. *See* baseless beliefs.  
 delusional, 36, 263-264, 338, 341, 763. *See also* delusional disorders.  
 immune system of, 43  
 immunity to metalogical argument, 645  
 immunity to revision, 264, 644, 650  
 privileging of, 39, 41  
 projective, 640-643. *See also under* projection.
- obstacles to overcoming, 642-647  
 psychology of, 645  
 without sufficient evidence, 360
- Berkeley, George, 118, 289, 294, 296, 401, 423  
 self-refutation argument, 117
- Berry's paradox, 61
- Berto, Francesco, 275
- Bethe, Hans, 579
- bifurcation of the human mind, 25, 357, 645
- Binet, Alfred, 737
- bivalence  
 accepted on metatheoretical level, 261, 720  
 principle of, 248, 719. *See also* excluded middle, principle of.
- Black, Max, 677
- Blackwell, Richard J., xxxvi
- Boas, Franz, 52
- Bochvar, Dmitri Anotolevich, 249, 266-267, 272, 274, 276, 721
- Bohm, David, 579  
 hidden variable proposal, 579
- Bohr, Niels, 569-570, 572-573, 575, 577, 581, 588, 590, 592, 594, 596-597, 599, 612, 624
- Born, Max, 573
- boundaries  
 epistemological, 320  
 of referentially forbidden, 301  
 of meaning, 224  
 of possible reference, 216. *See also under* horizons.  
 of reference, 224
- Boyle, Joseph, 104, 481
- Brentano, Franz, 65, 491
- bridge of rationality, 37 *See* rational bridge.  
*See also* rational bridge problem.
- Bridgman, P. W., 3, 39, 102, 126,

- 205, 209, 244, 257-259, 268, 283,  
294-296, 446, 496-497, 504-506,  
528-530, 532, 540, 772
- Brillouin, Louis Marcel, 527
- Brouwer, L. E. J., 381, 385-386, 688
- Bruner, J. S., 379, 737
- Buchanan, Scott, 701
- Burali-Forti paradox, 61
- Calder, Allan, 679
- Cantor, Georg, 686
- Cantor's paradox, 61
- Čapek, Milič, 582
- Capgras syndrome, 64
- Carnap, Rudolf, 78, 234, 266, 268,  
293-294, 296, 474, 491, 495, 500,  
506, 672, 730
- Cartesian *cogito*, 117-118, 160, 178,  
345, 666
- Cartesian coordinate system, 93, 96,  
213, 217, 220, 357, 548
- Castell, Lutz, xxix, xxxvi
- causal consequences  
as theory-based, 475
- causal field, 479
- causal network  
concept of, 477-479  
definition of, 478
- causality, 455-475, 481-484  
as functional dependency, 473-477  
conventional notion of, 460, 462-  
463, 465, 469, 472, 480, 483,  
485, 500  
experiential basis, 461-462  
primitiveness of, 472
- counterfactual  
in social sciences, 466
- de-projected concept of, 477
- de-projective understanding of,  
486-487
- doctrine of, 129
- evolution of concept of, 471-473
- framework-transcending, 480
- implicit shift from explanation to  
description, 469, 472, 476
- in general relativity  
replaced by principles of  
invariance, 472
- relinquished by quantum theory,  
472, 476
- relinquished by relativity physics,  
476
- temporal nature of, 462-463
- varieties of, 460
- cause and effect  
agency or tie between, 466-469
- counterfactual relation between,  
465-467
- necessary connection between,  
464-468, 477
- predictability, 468
- regularity, 468-471
- role of conceptual habit, 470
- role of psychological habit, 470,  
477, 481
- similarity relations between, 463-  
464, 478
- temporal priority of cause, 472, 486
- uniformity, 468-471
- cerebroses (Toulmin), 282
- certainty  
agent-centered, 352
- approaches to, 341
- evidential approach, 341, 343, 351
- intrinsic framework-relative, 352
- of de-projective analysis, 352
- pragmatically based, 346
- change  
of state, 446
- Chomsky, Noam, 52
- Chrysippus of Soli, 133
- Church, Alonzo, 249
- circular logic, 349. *See also* under  
circularity.
- circularity  
logical, 176

- of justification, 181
- presumptive, 178-180, 182-184, 348
- vicious, 195
- Clifford, William Kingdom, 290, 295, 360
  - ejects, 432-433
- clinical psychology, 32
- closed proofs, 366
- closed sets, 211, 213
- closed systems, 80, 335, 348, 439
  - ideological, 508
  - of thought, 754-755, 761
  - philosophical, 508
  - psychiatric, 508
- cogito*. *See* Cartesian *cogito*.
- cognitive differences
  - individual
    - resistance to acceptance of, 749
- cognitive disorders
  - according to Kant, 648-649
    - insania*, 648
    - vesania*, 648, 654
- cognitive impenetrability, 264
- cognitive disorders
  - projective delusion, 654. *See also* *under* projection.
- coherence, 355
- Coleridge, Samuel Taylor, 364
- Collingwood, R. G., 71-75, 191, 193
- Columbus, Christopher, 379
- combinatorial algebra, 303
- complementarity, 224, 589, 623. *See* principle of. *See also under* quantum theory.
- complementary descriptions, 224
- Compte, Auguste, 631
- computer vision, 223
- conceptual drive to over-reach, 414. *See* projection. *See also* horizons, human compulsion to transgress.
- conceptual limitation, 305
- conceptual pathology, 283, 649, 651
  - conceptual therapy, 283-284, 300, 647, 649
    - as applied to persons, 648, 663
  - conceptually unreachable, the, 391-416
    - and concept of horizon, 391-416
  - confirmation bias, xxvii, 745-750, 752
    - evidenced in philosophers, 751-752
  - congruence relation, 336
  - consciousness
    - as container of experience, 502, 509. *See* consciousness, myth of as a container.
    - myth of as a container, 502-504. *See also* myth of consciousness as a container.
    - in phenomenology, 81
  - consensus
    - popular, 9
  - consistency
    - defined in terms of possibility (Fitch), 135
  - constitutive structure, 219, 221, 223, 281, 312, 328, 330, 347, 352, 374, 443, 513, 591, 619, 664-665
    - definition of, 174
  - constraints
    - first-order, 210
    - metalogical, 210
    - parametric, 312
    - positive and negative, 281
    - second-order, 210
  - contradictions
    - realized, 135
  - coordinate system
    - concept of, 97
    - meaning in physics, 528
  - coordinate systems. *See under name of coordinate system.*
    - extrinsic, 548, 553
    - intrinsic, 548
    - mathematical conception
      - distinguished from physical

- frames of reference, 527
- tautological circularity of, 101
- coordinates
  - transformation of, 529. *See also* transformation of coordinates.
- Copenhagen interpretation, 135, 571-580
  - inappropriate naming of, 572-573, 583
- Copernican model, 284
- Copernicus, Nicolaus, 3, 666
- Coreth, Emerich, 731
- corrigeability, 344
- cosmology, 666
- counterintuitive
  - opposition to the, 421
- counterintuitiveness
  - of present work, 11
- covariance
  - application of concept of, 558
  - definition of, 557
  - general
    - definition of, 557
  - of physical laws, 452
- Crawford, Patricia, 202
- criteria
  - compelling in strong sense, 233
  - compelling in weak sense, 233
  - for judging sameness, 223
  - strongly non-arbitrary, 233-234
  - weakly non-arbitrary, 233
- criticism
  - immunity to, 253
  - internal, 253-254
- “Criticism of the Logistic Approach” (Minsky), 691
- critique of impure reason (the task of), 4, 300, 320, 485, 614, 625, 632, 663, 667, 670
- Critique of Impure Reason* (Bartlett), 4, 632, 635, 657, 665, 717
  - as negative science, 632-633
  - purpose of, x
- Critique of Pure Reason*, x, 4, 176, 194, 305, 307, 430, 447, 632-633, 635
- Cronus, Diodorus of Megara, 133
- Curry, Haskell B., 680
- curvature of a surface
  - intrinsic determination of, 549. *See also* Gaussian curvature.
- curvilinear coordinates. *See* Gaussian coordinates.
- cylindrical coordinate system, 217
- Dante (Dante Alighieri), 761
- Dawkins, Richard, 40
- de Broglie, Louis, 573, 579
- de Condillac, Étienne Bonnot, 51
- de Gaulle, Charles, 345
- De l’horizon de la doctrine humaine* (Leibniz), 303
- de Sitter, Willem, 542
- de Waehlens, Alphonse, xxxv
- deception
  - in Kant, 307
- decidability
  - in terms of recursion, 687
- Dedekind, Richard, 211-212
- deductive approach, 341-342
- Deecke, Lüder, 484
- definition
  - nature of, 69
  - real, 738
  - recursive, 686
  - stipulative, 738
  - varieties of, 738
- definition problem, 739-741
- delusion, 319-320, 333, 359
  - conceptual, 417
  - in Kant, 307
  - meaning of, 262
  - projective, 311, 359, 416. *See* projective delusion. *See also* under projection.
  - psychiatric definition of, 264

- psychiatric species of, 263
- delusional beliefs, 639, 763
  - anosognosia*, 639. *See also* anosognosia.
  - exemption from rules of rational admissibility, 639
- delusional disorders, 36, 81
- delusional thinking, 392, 495, 649.
  - See also* projection.
- delusions
  - anosognosia, 264
  - exemption from consistency, 264
  - exemption from rationality, 264
  - of control, 63
  - of grandeur, 64
  - of influence, 63
  - of meaningfulness, 264-265, 281, 297, 301, 650-651, 657
  - properties of, 651-654
  - of misidentification, 64, 263, 650
    - alien control, 263
    - Cotard delusion, 263
    - delusions of reference, 263
    - Fregoli delusion, 263
    - mirrored-self misidentification, 263
    - reduplicative paramnesia, 263
    - thought insertion, 263
    - unilateral neglect, 263
    - psychiatric sense, 263
  - of persecution, 64
  - projective, 650. *See also under* projection.
  - properties related to epistemology, 264
  - psychiatric theory of, 263
- deontic logic, 336, 636
- de-projection, 347
  - and tautological equivalence, 336-337
  - application to theories, positions, and concepts, 324
  - as transcendental heuristic, 327
  - cognitive skills required, 328, 335
  - corrective, 327, 329, 332, 366
  - descriptive, 327, 332, 365
  - descriptive task, 351
  - diagnostic, 327-328, 332, 365
  - eliminative, 327-328, 332, 365
  - empty of content, 332
  - focus not pragmatical, 324
  - logical structure of, 323, 332
    - tautologous, 330, 332, 336, 348
  - method of, 321-339, 657, 718.
    - as empty of content, 330
    - autarkic, 340
    - constructive purpose, 662
    - definitiveness of, 672
    - general applicability of, 626
    - heuristic nature of, 333, 351, 366
    - no new content introduced, 339
    - recurvature of, 419
    - reflexive nature of, 349
    - self-enclosure of, 508
    - tautologous nature of results, 331, 339, 673, 718
  - not algorithmic, 351
  - objectives of, 323
  - ontology of non-existent objects, 372-373
  - self-sufficiency of, 339. *See also* autarky.
  - stages of, 365
  - stipulative, 351, 365
  - theorem of, 331-332
- de-projective analysis, 359. *See also under* de-projection.
  - conceptual skills of, 441
  - four stages of, 327
  - psychological skills of, 441
  - self-validating, 366
- Descartes, René, 34, 506
  - evil genius, 127
- description
  - tautologous nature of, 330
- destructive testing, 184-185, 313, 316,



- 319, 328  
 in analysis of presuppositions, 313  
 in dynamic systems, 313
- determinism, 110, 475-479, 480-484  
 conventional notion of, 480  
 de-projective understanding of, 487  
 reduced to habit, 481
- Diagnostic and Statistical Manual of Mental Disorders (DSM)*, 264, 639
- diagonal procedure, 686-687
- dialetheic logics, 105, 275. *See also*  
 inconsistency-affirming logics.
- Dicke, Robert, 542
- dimension  
 as intrinsic property, 223  
 mathematical meaning of, 222
- Dirac, Paul Adrien Maurice, 573, 579
- discourse, domain of  
 in terms of game theory, 191
- discovery  
 in Einstein, 393-394  
 in mathematics, 381-388  
 in physics, 389-393  
 meaning of the term, 378  
 projection of, 400
- disorders of thought, 648. *See also*  
 thought disorders.
- dissonance, cognitive and emotional,  
 644
- distortion, in response to criticism,  
 646-647, 653
- distribution of mental health, 26, 736
- Divine Comedy* (Dante), 761
- Donceel, Joseph, 119-120
- Driesch, Hans, 491
- Drieschner, Michael, xxix, xxxvi
- Dumarsais, César Chesneau, 51
- economics  
 concept of autarky, 338
- egalitarianism, 49
- ego, the, 499  
 doctrine of, 129
- egocentric particulars, 61
- Einstein, 56, 66, 73, 97, 99, 174, 207-  
 218, 290, 318, 392-394, 426-427,  
 451, 472, 475, 480, 496, 523, 525,  
 528, 532-566, 572-577, 579, 583-  
 585, 588, 600, 603, 612, 616-621,  
 624, 626, 664, 666, 738, 771, 775-  
 776, 782, 785, 787, 789, 797  
 aim in physics, 534  
 appraisal of philosophy, 534  
 definition of science, 66  
 visual thinking, 56
- Einstein's mollusks, 548-551, 555,  
 664
- ejects, 290-291, 295
- electromagnetism, 212
- embedment, 214, 219, 223-224, 312,  
 513, 591, 619, 664  
 and identity, 214
- energy and mass  
 invariant functional  
 interdependency of, 207
- entailment, 167-171. *See also specific  
 types of.*  
 as a function of meaning, 167  
 contextual nature of, 167, 171  
 implicit, 253-255  
 metalogical. *See* metalogical  
 entailment.
- epistemological  
 abilities  
 individual variability of, 733,  
 738, 741, 747, 749, 750, 754,  
 756, 761
- admissibility, 359
- analysis  
 skill-based, 649
- boundaries, 184. *See also under*  
 horizons.
- intelligence, 673, 726-757  
 immunizing result of, 755  
 not known to be teachable, 761  
 potential role of talent, 761

- rejects transcendence, 763
- skills involved in, 734-736, 753-754
  - whether learnable, 761
- loop, 182-184, 188. *See also* circularity, presumptive.
- neutrality, 330, 522, 665
- rationality, 358, 360
- results
  - as intellectually "slippery", 37
  - understanding
    - pre-college level, 748-749
- epistemology
  - as a negative approach, 37
  - as a positive approach, 37
  - as content-based, 733
  - as skill-based 726-757, 733
  - circular nature of, 181
  - common definition of, 732
  - contributions to quantum theory, 624
  - contributions to relativity physics, 613, 624
  - influence upon physics, 612
  - internal limitations of, 183
  - limits of self-justification, 184
  - nature of, 37
  - skills of, 37. *See* epistemological intelligence.
  - transcendental, 131. *See under* transcendental argumentation.
- EPR (Einstein, Podolsky, Rosen) position, 574-577, 583
- equipollence, 336-339, 352
- erotetic intractability, 410, 412-413
  - projection of, 413
  - Rescher's view, 410-411
- erotetic logics, 636
- error
  - systemic in physics, 637
- error analysis and error theory
  - in physics, 637
- error avoidance
  - in computer science, 637
  - in mathematics, 637
  - in natural sciences, 637
- errors
  - due to imprecision of instruments, 637
  - propagation of, 637
    - with projective concepts, 638
  - statistical or random, 637
- Erwin, Erwin, 257
- Eslick, Leonard, 22
- Euclidean geometry, 355
  - parallel postulate, 342
- Evans-Pritchard, E. E., 652
- evidence, 360. *See also specific types of.*
  - admissibility of, 354
  - external, 351
  - ideal, 343
  - in law, 342
  - internal metalogical, 352
  - linked to proof, 342
  - root meaning of, 342
- evidential certainty, 343
- excluded middle
  - principle of, 355, 385, 719
- existentially inconsistent sentences, 345
- experience
  - as self-enclosed field, 335, 418
  - immanent, 344
  - phenomenologically recurved
    - space-time continuum, 81
  - pre-reflective, 501. *See also* pre-reflective experience.
  - relativistic constitution of, 130
  - self-enclosure of, 454, 508
  - self-neutral, 492, 502, 505-507. *See also* self, the, subjectless experience.
  - structural short-circuiting of, 694
- expert systems, 691, 695
- external

- ontological world, 425, 427, 429, 433
- others, 432
- perceptual space, 425, 428-429, 431, 432, 435, 447, 449-450, 518, 602
- questions (Carnap), 293
- reality
  - mind-relative (Berkeley), 289
- world
  - de-projected understanding of, 434-436
  - independence of, 426-427
  - problem of the, 416, 423
- facts
  - framework relativity of, 224
- Fadiman, Clifton, 355, 404, 763
- failure mode and effects analysis, 637
- faith, 40
  - religious, 40
- Far Side, The. *See* conceptually unreachable, the.
- Faraday, Michael, 212
- Farrell, Brian, 282
- fault tree analysis, 637
- Faust, 397
- fictitious objects, 202, 214, 216, 219, 298, 370, 373, 494
- field
  - concept in mathematics, 211, 213
  - concept in physics, 521
  - concept of, 174, 186, 211, 312, 514
  - concept of in physics, 515-516
    - modal nature of, 591
  - concept of is modal, 517, 521-522
  - effect, 517, 591
  - functional interdependency of, 516
  - mathematical
    - self-enclosure of, 418
  - of experience, 216
  - quantum-theoretical, 224, 418
  - referential. *See* referential fields.
    - scalar. *See* scalar field.
    - theory, 211, 280, 517
    - vector. *See* vector field.
- Fitch, F. B., xxxv, 75-78, 134, 405-409, 413, 419, 663, 704, 769, 773, 777, 802
- Fitch's Paradox of Unknowability, 405. *See also* Fitch's Theorem.
- Fitch's Theorem, 405-408
  - projective nature of, 407
  - referential preconditions of, 407
  - role of time, 409
- formal limitation
  - theorems of, 680-681
- formal ontology. *See* ontology, formal.
- formalism
  - in mathematics, 387
- formalization
  - advantages, 676-678
  - deductive, 691
  - disadvantages, 676-680
  - internal limitations of, 674-676, 680, 690
  - non-deductive, 691-692
  - of human experience, 694
  - of metalogic of reference. *See under* metalogic of reference.
  - See also under* MoR.
  - possible inescapable limitation of, 694
  - of ordinary reasoning 684-690
  - total, 676-679, 682-683
    - commonalities with metalogic of reference, 683
    - impossibility of, 684, 695
- formalized systems
  - impossibility of total reflexivity of, 685, 689, 695
- frames of reference
  - as dynamic referential systems, 513, 638
  - as systems of coordination, 96,

- 209, 327, 525  
cannot be reified, 513  
compatibility of, 220, 230  
  definition of, 219  
complementarity, 221, 223-224, 526  
  definition of, 223  
concept of, 512  
coordinative nature of, 512  
definition of, 208  
dimensional incompatibility of, 222-223, 230, 608  
  in quantum theory, 608  
in physics, 527. *See also* reference frames.  
isomorphism, 221, 526  
meta-frameworks, 220, 223  
spatial, 424  
translatability between, 223, 528  
framework relativity, 99, 162, 201, 203-224, 236, 255, 262, 264, 278, 280, 293-294, 296, 323, 326, 331, 344, 351, 366, 373-374, 386-388, 390, 418-419, 441, 486, 508-509, 511-513, 522-524, 536-537, 562-564, 566, 568, 574-575, 588, 590, 592, 616-620, 623-624, 662-664, 673  
linguistic, 294  
of experience, 435  
of extendability, 416  
of identifiability, 416  
of meaning, 416  
of reality, 416  
opposed to relativism, 205  
two principal routes to the study of, 614-615  
frameworks of reference. *See* frames of reference.  
framework-transgressing concepts, 130. *See* projection.  
Frank, Philipp, 539  
free choice  
  arguments defending, 112  
  dogma of, 482  
  legal role of belief in, 482  
  rejection as self-falsifying, 128  
free will, 477-484  
  acts of choice, 483  
  acts of volition, 483  
  belief in, 482  
    by evolutionary biologists, 482  
    by philosophers, 482  
  conventional notion of, 483  
  de-projected rejection of, 487-488  
  neurological research, 484  
  projective nature of, 482-483  
  role in ethics of belief in, 482  
  role in psychiatry of belief in, 482  
  role in religions of belief in, 482  
Freese, Ray, xxxvi  
Frege, Gottlob, 53, 59-60, 381-382, 769, 777-799, 804  
Freud, Sigmund, 44, 285-287, 642, 778  
fringes  
  doctrine of, 728  
function  
  concept of in mathematics, 474  
functional dependency  
  definition of, 474  
functional interdependency, 207, 475-476, 515-517, 521-522, 591, 598-599, 606, 608, 664  
future  
  framework-independent, 445  
  independently existing, 446  
Gagné, R. M., 379  
Gal, David, 38  
Galilean reference frames. *See* inertial reference frames.  
Galilean transformation, 538-539, 543, 545-557  
Galileo (di Vincenzo Bonaulti de Galilei), 666, 673

- game theory, 191, 355
- games  
   rule-based, 191-192
- Gardner, Howard 737
- Gauss, Johann Karl Friedrich, 547-548  
   theory of surfaces, 547
- Gaussian coordinates, 548-550, 555-556
- Gaussian curvature, 557
- general principle of relativity, 545, 550
- general relativity, 315-316. *See also* general theory of relativity.
- general systems theory, 61, 175, 772
- general theory of relativity, 97-98, 208, 472, 523, 528, 532-533, 537, 539, 546-547, 551-552, 556-566, 664  
   covariance of physical laws, 557  
   departure from Euclidean geometry, 547  
   invariance of physical laws, 557  
   metric tensor, 549  
   objective of, 545, 553  
   relinquishes absolute space and time, 451  
   role of tensors, 557, 559
- generalized coordinates. *See* Gaussian coordinates.
- genidentity, 352
- geometry  
   Euclidean, 206  
   Lobashevskian, 206  
   Riemannian, 206
- germinal duplication  
   of system of reference, 214
- Getzels, J. W., 379
- Goddard, Leonard, 273, 260, 272, 276, 279  
   matrices for connectives, 274
- Gödel, Kurt, 381, 680  
   incompleteness theorems, 681
- Goleman, Daniel, 737
- GPS coordinates, 186, 315-316
- Graham, Paul, 53
- grammairiens-philosophes*, 51
- gravitation, 463  
   and curvature of space, 207  
   in general relativity, 476  
   theory of, 212
- gravitational field, 462, 516-517, 538, 542, 546-550, 559, 618  
   variability of speed of light in, 542
- gravitational potential energy, 517
- gravitational potential function, 517
- Greene, Marjorie, xxxv
- Grelling's paradox, 61
- Grier, Michelle, 640
- Griffiths, A. Phillips, 596, 610
- Grisez, Germain, 104, 481
- Gruber, H. E., 379
- Grynypas, Jérôme, 80
- Guilford, J. P., 379, 737
- Gurwitsch, Aron, 74, 77
- Hacker, P. M. S., 42
- Hadamard, Jacques, 56
- Halldén, Sören, 249, 267-268, 270-274, 276, 279  
   matrices for basic connectives, 269-270
- Halmos, Paul, 674
- Hanson, Norwood Russell, 579
- Harris, Sam, 481-482
- Hartmann, Klaus, 349
- Heelan, Patrick, 575, 729
- Hegel, Georg Wilhelm Friedrich, 349, 417, 419
- Heidegger, Martin, 389-391
- Heisenberg, Werner, 568, 572-573, 575, 579, 581, 583-584, 594-596, 600, 610, 612, 624  
   *potentia*, 595
- Heisenberg's microscope, 568
- Henkin, Leon Albert 680

- Henri, Victor, 737
- Herder, Johann Gottfried, 52
- heuristic, 353, 363
  - rationale for, 317
  - axiomatized, 696, 703, 705
  - method, 315-317
    - definition of, 317
    - advantages of, 318
    - in Einstein, 318
    - non-algorithmic, 320
  - principles. *See* metalogical heuristic principles.
- Heyting, Arend, 382-384, 688
- Higgs field, 374, 516
- Hilbert, David, 386-387
- Hintikka, Jaakko, 253-346, 348
- Hjelmslev, Louis, 173
- Hocutt, Max O., 652
- Horace, Horatius Flaccus, 3
- horizons, 300-320, 333, 336-337, 347, 356, 361, 394, 405, 417, 420, 430, 439
  - as modal limit, 396
  - concept of, 142, 203, 319
    - in Coreth, 731
    - in Heelan, 729
    - in Husserl, 728
    - in Kuhn, 729
    - in Lonergan, 730-731
    - in other disciplines, 727
    - in transcendental Thomism, 730-731
  - in work of other philosophers, 727, 728-729, 730-731
- definition of, 303, 310
- delusion-inducing character of, 325, 396
- human compulsion to transgress, 309, 388, 402, 413, 420, 429, 439, 458, 487, 512, 576, 592, 640-641, 649, 660-661, 694, 758
- psychological nature of, 640-642
- psychological obstacles to
  - overcoming, 642-647
- intuitive properties of, 311
- Kant's concept of, 305-307
  - transcendental, 122
  - transgressions of, 307-308
- Leibniz's concept of, 303-304
- limits of, 305
- origins of term, 303
- phenomenological concept of, 728-731
- philosophical compulsion to transgress, 641
- physical, 303, 305, 396, 415
- properties of, 319
- reactive boundaries of, 311, 319, 325, 419, 439, 522, 688. *See also* horizons, symptomatologically detectable.
- symptomatologically detectable, 311, 319, 397
- transgression of, 203, 265
- transgressions of. 203, 265. *See* horizons, human compulsion to transgress. *See also* under projection.
- transgressions of in Kant, 310
- unboundedness of, 303, 305, 396
- horizontal theories, 75
- human reasoning
  - algorithmic representation of, 691, 695
  - formalizability of, 690-692, 695
- Hume, David, 43, 109, 234, 463, 465-471, 477, 481, 534, 782-783, 790, 801
- definitions of causality, 463, 465
- Husserl, Edmund, xxxiii, xxxiv, 22, 65, 73-74, 76, 130, 194, 336, 339, 352, 447, 449, 491, 672, 699-700, 728-729, 743, 766, 782-783, 785
- Hutchins, Robert M., xxxv
- Huxley, Aldous, xxx
- hyperbolic hypothesis (Bridgman),

- 244, 283  
 hypothetico-deductive method, 186, 341
- idealism, 416, 436, 439-440, 489, 503, 666  
 event-idealism, 436, 438  
 meaninglessness of (Carnap), 293  
 metalogically self-undermining.  
   *See* projection, of idealism.  
 motivation for, 440  
 projection of, 439-440. *See also*  
   projection, of idealism.  
 views that oppose, 110
- identifiability, 312  
 framework relativity of, 203, 215  
 maximally general sense of, 204  
 object-related, 188  
 system-related, 188
- identification, 95-96, 101, 140, 186-190, 209, 224, 239, 281, 301, 531.  
*See also* reference, as identification.  
 abstract, 186  
 according to rule, 186  
 as activity or process, 187  
 as unique specification, 186  
 bivalence and trivalence, 189  
 component of satisfaction, 239  
 concept of, 69  
 conditions of, 188. *See also*  
   presuppositions, of identification.  
   not linguistic, 216  
 descriptive component of, 239  
 epistemologically naive view of, 187  
 erroneous, 264, 640  
 framework relativity of, 239  
 general, 186  
 intrinsic, 545-548  
 meaning of, 187, 208  
 metalogic of, 189  
 metalogical conditions of, 190, 215  
   structural/systemic nature of, 189  
 metalogical properties of, 241  
 preconditions of, 139, 213  
 system-based, 517  
 vague, 186  
 varieties of, 186, 219
- identification frameworks  
 epiphenomenal properties of, 213-214, 216
- identifying reference, 190, 199, 208-209, 216, 225, 231, 237-241, 260, 277-279, 296, 316, 331, 357, 372, 398, 418, 431, 442-443, 457-458, 463, 531, 548, 553, 562, 565, 569, 582, 663, 685-686, 688, 698-700, 702, 707-708, 710, 713-714, 716.  
 and meaning  
   form bonded pair, 277-279, 281, 312  
 as ternary relation, 237-239, 518  
 component of satisfaction, 208  
 definition of, 239  
 descriptive component of, 208  
 meaning of, 187, 208  
 metacondition of, 140  
 physical, 540. *See also* physical identifying reference.  
 physical variety of  
   definition of, 526-527  
 temporal basis of, 443, 448, 453  
 temporal nature of, 685  
 to material bodies, 450-451  
 to physical objects, 455-457  
 to physical phenomena, 537
- identity, 102  
 concept of, 69  
 identity conditions, 701  
 framework relativity of, 201, 213, 231, 562  
 function of parametric constraints, 281

- inseparable from reference system, 127
- of objects, 188, 201, 208, 223, 562
  - framework relativity of, 201
  - personal. *See* personal identity.
  - system-based, 517
- ideological self-encapsulation, 646-647, 653
- idéologues*, 51
- ideology
  - immunity to falsification, 643
- ignorance
  - unlimitedness of, 400
  - de-projective understanding of, 400
- illusions
  - cognitive, 262
  - sensory, 262
- immunity
  - to criticism, 345
  - to challenge, 343, 352
- immunization of belief
  - psychological principles governing, 645-646
- imperviousness to criticism, 31, 646, 742, 745
- implication
  - material, 166
- implicit commitments, 117-118
  - of coherent assertions, 117
- Inaugural Dissertation* (Kant), 451
- incoherence, 356
  - projective, 355
  - self-defeating, 356
- incorrigibility, 343-344, 352
- indefiniteness, 209, 239
- indeterminacy, 288, 623. *See also* uncertainty principle.
- indexical signs, 61
- indirect proof, 333-334. *See also* *reductio ad absurdum*.
- individual cognitive differences
  - resistance to acceptance of, 749
- indubitability, 343, 352
- induction
  - mathematical, 686
  - principle of, 470
  - problem of, 470
    - application to causality, 471
- inductive reasoning, 469
- inertial reference frames, 207, 538, 542-547, 550, 553, 557-560, 564, 618
- inexpressible
  - expressing the, 403-404
- infallibility, 343, 352
- infinite
  - concept of in mathematics, 386-387
  - projective nature of, 387
- Institut Catholique de Louvain, xxxv
- intelligence, 356
  - emotional, 737
  - epistemological. *See* epistemological intelligence.
  - general
    - improvement through teaching, 759
    - moral, 737, 759
    - multi-dimensional conception of, 737
    - social, 737
- intelligibility, 347, 356, 359
  - meaning of, 356
- intentionality, 65, 491, 729, 793, 804
- intermetamorphosis, 64
- internal limitations
  - of human understanding, 320
- internal questions (Carnap), 293
- intrinsic differential geometry, 552
- intuition
  - benefits and shortcomings, 422
  - definition of, 421-422
  - mathematical, 684
  - of two-oneness (Brouwer), 385, 688
  - primitive nature of, 422



- intuitionism, 110, 688  
 in mathematics, 381-387, 393
- invariance, 71, 73, 80, 82, 131, 189, 280, 475  
 Galilean, 557  
 in relativity physics, 475  
 meaning of, 557  
 of speed of light, 207
- invariance theory. *See Invariantentheorie.*
- invariant conditions of discourse (Passmore), 112-114, 117-118
- Invariantentheorie* (invariance theory), 207  
 as more appropriate name for relativity physics, 207, 535
- invention  
 in Einstein, 393-394  
 in mathematics, 381-388  
 in physics, 389-393  
 meaning of the term, 378
- irrational persistence, 30
- irrationality of  $\pi$ , 633
- irrefutability  
 framework-relative meaning of, 350
- Isaye, Gaston, 112, 119-125, 130, 194, 602, 731, 783, 789  
 approach is activity-based, 125
- isomorphism  
 metaphorical, 221
- Iverson, Kenneth E., 53
- Jackson, P. W., 379
- Jacobi, F. H., 430
- James, William, 291, 296, 728
- Jammer, Max, 461
- Jeffries, C., 579
- Johnson, Samuel, 47, 370, 374, 504
- Johnstone, Henry W., Jr., xxxv, 28, 61, 105-114, 123, 125, 176, 359, 783
- justification  
 epistemological, 195
- Kant, xxxiii, 4-5, 27, 43, 51, 78, 97, 99, 119, 122, 126, 128, 130-131, 134, 176-179, 181, 194-196, 256, 279, 300, 305-310, 317, 319-321, 324, 326, 381, 386, 402, 413, 419, 429-430, 447, 451-452, 458, 534, 576, 632-637, 640, 642, 648-649, 653-654, 658, 729, 774, 779, 784, 789-791, 800  
 absolute world-space, 451  
 compulsion to transgress horizon, 307-308, 310, 319, 326, 386, 413, 429  
 concept of horizon, 310  
 boundaries of, 305-308  
 deception in, 307-308, 324  
 delusion in, 307-308, 324  
 Forms of Sensibility, 447  
 on conviction, 27
- Kaufman, Walter, 669
- King Hiero II of Syracuse, 317
- King James Bible*, 497
- Kleene, Stephen Cole, 268, 680
- Klein  
 bottle, 80  
 surface, 199
- Kneale, William and Mary, 147-148
- Kneser, Adolf, 535
- knowability principle, 408
- knowledge  
 as coordination, 97-98, 537, 608  
 finitude of, 397-399  
 projection of, 398, 408  
 putative meaning of, 408-409  
 incompleteness of, 399, 419  
 projection of, 399  
 transcendental theory of, 97
- Kornhuber, Hans Helmut, 484
- Kripke, Saul, 60
- Kuhn, Helmut, 729
- Kuhn, D., 748

- Ladrière, Jean, xxxv, 124, 680-690, 783, 785
- Lambert, xxxiii, 256, 633-635, 637, 786, 791, 801
- Langford, C. H., 147
- language
- feature-placing, 700
  - logically posterior to preconditions of identifying reference, 278
  - natural, 51-52
  - ordinary, 49-50, 52, 56, 58, 92, 144, 172, 187, 249, 282, 510, 617
  - process-oriented, 700
  - subject-predicate structure of, 491
- language-games, 700
- Laszlo, Ervin, 175
- Leibniz, Gottfried Wilhelm Freiherr von, 137, 303, 309
- concept of horizon, 310
  - finitude of expressible truths, 304
  - use of combinatorial algebra, 303
- Leibniz boundaries, 305-308, 312, 336, 396, 399, 409, 415-416, 419, 425, 431, 433, 435, 439, 449, 453-454, 470, 478, 494, 518, 688
- definition of, 305
- Lemmon, E. J., 149, 706
- length contraction, 542
- Leopardi, Giacomo, 40
- level of maximum theoretical
- generality, *See* maximum theoretical generality, level of.
- Levi-Civita, Tullio, 547, 552
- Lewin, Kurt, 352
- Lewis, C. I., 144, 146-147, 149-151, 153, 155, 496-497
- Lewis, David, 138, 464
- Libet, Benjamin, 484
- Library of Living Philosophers, 23
- Lichtenberg, Georg Christoph, 506
- light
- constancy of speed of
    - as a supposition, 541-544, 546, 559
  - contrary properties of, 223
  - particle or wave phenomenon, 221
  - variability of speed of, 542
- limitations
- conceptual, 320
  - neurological, 320
  - of language, 320
  - of reference, 82
- Limitations internes des formalismes, Les*, (Ladrière), 680
- limitative boundary conditions, 310.
- See* horizons.
- linguistic
- conditions of discourse, 215
  - relativity hypothesis, 52-53, 61, 747
- List, Peter, xxxvi
- Locke, John, 87, 597, 614
- logic
- calculus of, 633
  - consistency of, 693
  - of meaning, 333
  - of meaninglessness, 265-266, 276, 281, 333
  - of nonsense, 267, 270. *See also* logic, of meaninglessness.
  - logic of structure, 81, 173-174, 728
  - medieval, 336
  - systems of
    - inconsistency-affirming, 105
    - inconsistency-tolerant, 105
    - many-valued, 105
    - three-valued, 135, 168-169, 189, 249, 288, 720-721
- Logic of Nonsense, The* (Halldén), 268
- Logic of Significance and Context, The* (Goddard & Routley), 272
- logical
- completeness
    - question of need for, 693

- consistency  
     question of desirability of, 692-693, 695, 726  
 criteria  
     arbitrariness of, 232-233, 235  
 paradoxes, 260  
 positivism/verificationism, 468  
 rules  
     as conventions, 232  
 standing  
     meaning of, 672. *See also*  
     metalogic of reference,  
     logical standing of.  
 systems  
     completeness of, 692  
*Logical Structure of the World, The*  
 (Carnap), 235  
 logicism  
     in mathematics, 382, 386  
 Loneragan, Bernard, 730-731  
 Lorentz, Hendrik Anton, 218, 533,  
 543-544, 573, 612  
 Lorentz transformation, 218-219, 539-  
 545, 547, 556-557  
     formulation of, 544  
 Lorenzen, Paul, 210, 344, 699, 706  
 Loux, Michael J., 144  
 Löwenheim, Leopold, 680  
 Łukasiewicz, Jan, 78, 133, 146, 149-  
 150, 153-155, 245, 249, 267, 788,  
 790  
  
 MacCoun, Robert J., 41  
 Mach, 491, 534  
 Mackie, J. L., 61, 104, 112, 115-119,  
 125, 215, 260, 397-398, 405, 473-  
 474, 479, 788  
 Madame Bovary, 374  
 Magueijo, João, 542  
 many-valued logics, 636  
 Marbach, Eduard, 65  
 Marcel, Gabriel, xxix  
 Marcuse, Herbert, xxxv  
  
 Maréchal, Joseph, 119-120, 731, 788  
 Markowitz, Sally, xxxvi  
 Martin, John Levi, 461, 466, 473  
 Martin, R. M., xxxiii  
 mass-energy equivalence, 475  
 material implication, 166  
*Mathematical Principles of Natural  
 Philosophy and the System of the  
 World* (Newton), 449, 612  
 mathematics  
     discovery in, 376-388, 390  
     epistemology of, 376  
     invention in, 390  
     results of  
         framework relativity of, 386  
         ontological status of, 390  
     solutions in  
         discovered or invented, 128,  
         376-388  
 matrix  
     for biconditional, 250  
     for conjunction, 249  
     for disjunction, 250  
     for implication, 250  
     for negation, 249, 250  
 Maugham, W. Somerset, 424, 503,  
 654  
 Mauthner, Fritz, 57-59, 282, 771, 789,  
 803  
 maximum theoretical generality, level  
 of, 6, 68, 70, 72-73, 75-77, 81, 94,  
 132, 154, 217, 225, 280, 335, 448,  
 512, 558, 603, 615, 657, 663-664,  
 674-675, 690, 695, 705, 717  
     self-enclosure of, 418  
 Max-Planck-Gesellschaft, xxxvi  
 Max-Planck-Institut zur Erforschung  
 der Lebensbedingungen der  
 wissenschaftlich-technischen Welt,  
 xxix, xxxvi  
 Maxwell, Clerk, 174  
 Mayer, John D., 737  
 meaning. *See also* referential

consistency, as criterion of  
 affective, 79  
 boundaries of, 281  
 boundaries of possible, 306. *See also* horizons.  
 catastrophic failure of, 638  
 cognitive, 79  
 concept of, 49, 202  
 criteria of, 231, 278  
   arbitrariness of, 235  
   as external standards, 232, 235  
   decline of philosophical interest  
   in, 257  
   failure of, 232, 257  
   generally do not compel assent,  
   235  
   inescapable criterion of, 232  
   not strongly compelling, 257  
   resistance to, 235  
   self-application of, 235  
   strongly non-arbitrary, 236  
   variety of, 234  
 delusional, 262. *See also* meaning,  
   problem of putative.  
 domain of  
   negatively defined, 320  
 domain of possible  
   negatively defined, 301  
 framework-relative nature of, 209  
 framework relativity of, 203, 231,  
   265  
 general theory of, 202  
 limits of, 225  
 linguistic, 228, 230, 278  
   dimensions of, 230  
 metalogic of, 225-255  
 metalogically self-undermining,  
   203  
 necessary condition of, 225, 248,  
   251  
 need for many-valued logic, 288  
 openness of definition of, 231  
 possible, 4, 250, 263. *See also*

*under* horizons.  
 pre-analytically accepted, 262, 299  
 preconditions of possible, 263  
 problem of putative, 259-260, 311,  
   413  
 referential consistency  
   as criterion of, 231-255, 333,  
   355, 518  
   as necessary condition of, 231  
 requires identifiability, 231  
 self-destruction on level of  
   possible, 259  
 subsequent to analysis, 299  
 sufficient condition of, 226  
 theories of, 79, 225, 228  
   activity-based, 227  
   concept-based, 227  
   contextual-relational, 227  
   dictionary-based, 227  
   information-based, 228, 230  
   pragmatical, 227  
   referential, 227  
   rule-based, 227  
   varieties of, 226-230, 242, 288  
 truth-functional definition of, 266  
 verifiability as criterion of, 257,  
   271  
 meaningfulness  
   boundaries of possible, 361  
   delusion of, 262, 264. *See also*  
   meaning, problem of putative.  
   thresholds of, 281, 300, 317, 320  
 meaninglessness  
   as defined by Goddard and  
   Routley, 274  
   as defined by Halldén, 268-271  
   as unintelligibility, 356  
   post-analytical, 311  
   projective, 267  
   relativist view of, 258  
   symbol for, 245  
   truth-functional, 276  
 measurement-based perturbation, 568-

- 571
- Meaning of Meaning, The* (Ogden & Richards), 226
- Meiklejohn, J. M. D., 307
- Meinong, Alexius, 374
- Mellor, W. W., 259
- memory, 261, 442
- Mendelssohn, Moses, 635
- mental health
- distribution of, 26, 736
- Mermin, N. David, 573
- meta-frame indicators, 240
- metalogic
- definition of, 78, 225
- metalogic of identification, 69, 189
- metalogic of reference, xxix, 76, 217, 279, 288, 296, 301, 303, 326, 341, 347, 349, 356, 360, 388, 418, 614, 616, 657
- affinity with quantum theory, 611
  - affinity with relativity physics, 536, 545, 564
  - applications of, 360, 363-4. *See under specific philosophical problems.*
  - applications of to physics, 625
  - as a task of pathology, 717. *See also* conceptual pathology.
  - as axiomatized heuristic, 696. *See also* heuristic, axiomatized.
  - as negative science, 639. *See under* negative science.
  - as theory of error analysis and correction, 636-639
  - as transcendental, 79
  - central role of invariance in, 556-559
  - commonalities with total
    - formalization, 683
  - concept of objectivity in, 603
  - confirmation of results by quantum theory, 615-617
  - confirmation of results by relativity
    - physics, 615-617
  - confirms results of relativity
    - physics, 564
  - confirms results of quantum theory, 619-623, 639
  - contrast with relativity physics, 563-564
  - contributions to epistemology, 625
  - covariance of results, 562
  - defined in terms of heuristic, 300
  - definition of, 225
  - formal structure of 669-720
  - formalization of, 237-252, 260. *See also under* MoR.
    - possible inescapable limitation of, 694
  - general applicability, 626
  - general covariance of results, 557-559
  - general ontology of objects, 373-374
  - heuristic nature of, 365, 563, 695
    - formal structure of, 695-697
  - logical standing of, 672-674
  - logical structure of
    - potentially not deductively formalizable, 690, 695
  - main tasks of, 317, 631-632
  - maximally fundamental nature of, 68, 635, 663-664
  - maximally general nature of, 210
  - meaning of, 65, 68-69, 188
  - methodological principles of, 562
  - modal level of, 565
  - need for three-valued logic, 719
  - negative results of
    - positive value of 652-668
  - negative task of, 300
  - not traditional transcendental
    - argumentation, 131
  - objective of, 208, 268, 281, 609, 636, 694, 717
    - description of invariants of

- framework relativity, 536
- ontologically open, 661, 718
- philosophical approach of, 356
- pragmatical application of, 694
- relation to eliminative psychology, 641, 658
- results reached, 560
- revolutionary nature of, 667-668
- role of the observer, 518-522
- self-enclosure of, 508-509
- self-referential capacity of, 685
- shift from extrinsic to intrinsic reference, 551, 563, 619, 664
- transcendental foundation of, 210
- understanding of reality, 661-662
- Metalogic of Reference* (Bartlett), xxix, 5
- metalogical boundaries, 336. *See under* horizons. *See also under* metalogical horizons.
- metalogical entailment, 243, 252-255, 331, 705, 710
  - symbol for, 238, 243, 705, 710
- metalogical heuristic principles, 251
- metalogical horizons, 4, 204, 310-312, 319-321, 325, 333, 359, 429, 436, 440, 465, 471, 485, 493, 501, 508, 512, 564, 582-593, 608-609, 610, 615, 617, 632, 636, 638, 641, 657, 696-697, 717, 762. *See also under* horizons.
- of referential fields, 522. *See also under* referential fields.
- transgressions of, 336. *See also* horizons, human compulsion to transgress.
- metalogical projections. *See under* projection.
- metalogical rejection, 245, 247, 359
  - operator, 246-247, 250
- metalogical self-reference, 104, 126-130, 298
  - distinguished from pragmatical self-reference, 126
- metalogical self-referential inconsistency, 85, 201, 242, 247, 250, 320, 347
- metalogical self-validation, 247. *See also* self-validation.
- metalogically projective contradictories, 247
- metalogically projective contraries, 247
- metamathematical reasoning, 681-682, 694
- metaphysical self, 128, 130
- metatheory
  - derivability sign, 238
  - rejection operator, 238
- method of natural deduction, 342
- Michelson-Morley interferometry experiment, 542
- Mill, J. S., 52
- Minsky, Marvin Lee, 691-694, 726
- misidentification
  - delusions of, 263
- Möbius strip, 48, 80, 199
- modal logic, 144-146
  - appropriate for metalogic of reference, 151-154
  - basic (Łukasiewicz), 150
  - iterated modal operators, 147, 150, 152-153, 155
    - non-reducible, 153
  - levels of modality, 152
- M, 151
- need for more than two truth values, 153-154
- obstacles to philosophical use of, 145-146
- reduction rules, 150, 152
- S1, 147, 153
- S1-S5, 147-148
- S4, 151-152
- strong reduction principle, 147
- T, 151, 153

- weak reduction principle, 147
- modal operators
  - multiple, 146
- modalities
  - higher-order, 146-151, 155
- Moh, Shaw-kwei, 249, 267-268, 276, 721
- Moleski, Martin X., 120, 123
- mollusks
  - Einstein's. *See* Einstein's mollusks.
- Moore, G. E., 117
  - defense of common sense, 111
- MoR (axiomatized heuristic of metalogic of reference) 691-719, 697
  - as axiomatized heuristic, 703
  - concept of descriptions in, 702, 712
  - concept of identification in, 702
  - concept of identifying reference in, 698
  - concept of particulars in, 698, 700-702, 707, 709, 711-713, 715-716, 719, 721
  - deductive representation 696-710, 713-717, 719
  - definition of particulars in, 699
  - elements of deductive theory in, 703-708
  - heuristic inference in, 725
  - identification frameworks, 699, 702, 714-715, 722
    - definition of, 698
  - identifying reference in
    - definition of, 707
    - logically primitive in, 707-708
  - law of excluded middle
    - modal form of, 709
  - law of non-contradiction
    - modal form of, 709
  - ontological conclusion of, 719
  - ontological neutrality of, 718
  - possibility
    - logically primitive in, 707-708
  - projection
    - formal expression of, 720, 722
  - rule
    - of double negation for possibility, 709
    - of possibility introduction, 709
    - of referential counter-exemplification, 710
    - self-validation in, 706
    - use of arbitrary names in, 706
- moral bridge, 758-759
  - problem, 25, 759
- moral intelligence, 737, 759
- moral reasoning, 25, 759
- Morgan, Conwy Lloyd, 291, 296
- Morris, Charles, 88
- Morrison, Roy D., 309
- Mostowski, Andrzej, 249
- motion
  - framework relativity of, 546
- Muck, Otto, 123
- multi-culturalism, 20
- Murray, Charles, 13
- myside bias. *See* confirmation bias.
- myth of consciousness as a container, 502-503, 505, 507, 509
- narcissism
  - clinical, 743-744, 754
  - intellectual, 30, 35, 81, 743, 755
  - personal, 35
- Nārgājuna, 403
- Natanson, Maurice, xxxv
- Natorp, Paul, 491
- Natur und Kunst* (Goethe), 321
- nature, laws of
  - invariant and universal, 208
- necessitation
  - rule of, 151
- necessity
  - framework-relative understanding of, 201
  - framework-relativity of, 200

- in terms of parameters of constraint, 200
  - in terms of possibility, 143
  - modal concept of, 132
  - theory of, 142-143
- negation
  - kinds of, 245
  - metalogical, 245
  - traditional, 245, 250
- negative science, xxxiii, 4, 256, 279, 301, 320, 333, 459, 485, 489, 507, 604, 607, 613-617, 619, 621, 623-624, 626, 631-639, 641, 643, 657-658, 662, 667, 670
  - non-Kantian meaning in this work, 633
  - Kant's conception of, 635
  - results of, 614
- Nelson, Leonard, 176-178, 181-184, 188, 195-196, 198, 210, 348
- Neues Organon* (Lambert), 633
- neural networks, 691
- Newell, A., 317
- Newton, Isaac, 212, 389-391, 449, 454, 516, 538-539, 612
  - absolute time, 128, 446, 455, 457, 538, 540, 560, 564, 659
  - absolute space, 128, 449-451, 454-457, 538, 560, 564, 618, 659
- Newton's
  - gravitational constant, 475
  - second law, 475
- Newtonian
  - mechanics, 390
  - physics, 536-537
    - absolute length, 540
    - absolute time, 540
- Nietzsche, Friedrich, 491, 668
- non-contradiction
  - modal law of, 136
  - principle of, 121, 355
- non-distributive lattices, 234
- non-Euclidean geometry, 350, 553, 633
  - non-existent objects, 369-371, 374
  - non-inertial reference frames, 538, 545-546, 557-559
  - non-omniscience. *See* knowledge, finitude of.
  - non-ordinal theories, 75, 77, 217, 335, 418, 508, 603, 663, 674-675
  - nonsense, 260. *See also* meaninglessness.
  - Norman, J., 275
  - Northrup, F. S. C., 569
  - noumena, 307-308, 321, 402, 429-430, 640
  - Novum Organon* (Bacon), 633
  - number
    - concept of, 688
    - theory, 103
- object
  - general definition of, 298
  - of reference
    - framework-relative reality of, 224
    - presumption of single underlying, 221
- objectivity
  - framework-relative, 602
  - of formal systems, 602
  - of invariant natural principles, 602
  - in quantum theory, 427, 602-609, 623
  - meaning of, 206
  - metalogical, 603
  - of complementarity, 603, 606, 623
  - of formal systems, 606, 623
  - of invariant natural principles, 606, 623
  - of perspectively given information, 602, 606, 623
  - of rationality, 602, 606, 623
  - of science



- arguments defending, 112
- principle of, 121
- projective notion of, 603
- self-referential inconsistency of
  - denial of, 207
- objects
  - framework relativity of, 373
- objects of reference, 221
  - identity of, 223
  - incompletely given, 415
  - putative, 327, 443, 720
- observational frames of reference, 548
  - definition of, 528
  - inertial, 538
  - non-inertial, 538
  - transformation between, 538
- observer, the
  - ability to retain information about
    - past states, 519
  - commonsense notion of, 520
  - concept of, 519-521
  - non-human, 519
  - role of in metalogic of reference,
    - 518-522.
- ontological openness, xxxiii
- ontology, 215, 369, 370-375
  - closed, 718
  - formal, 215
    - definition of, 189
  - framework relativity of, 216, 219,
    - 280, 374-375, 661
  - open, 718
- open source (open access) publishing,
  - 25
- operational self-refutation (Mackie),
  - 115-119, 397
- operationalism, 259, 268, 294-295
  - as criterion of meaning, 257
  - operationally meaningless
    - statements, 249
- operative logic, 344, 346-347, 352,
  - 706
- Oppenheimer, J. Robert, 579
- ORCID identifiers, 186
- ordinal theories, 75-76, 663
- ostensive reference, 63, 87-88, 92,
  - 277
- Othello, 374
- other minds. *See* projection, of other
  - minds.
  - continuity of, 435-436
  - de-projected understanding of, 434-
    - 436
  - regularity of, 435-436
- Outlines of Pyrrhonism* (Sextus
  - Empiricus), 401
- Pais, Abraham, 528
- Palmer, Humphrey, 176, 178-185,
  - 188, 195-196, 198, 210, 348
- paraconsistent logics, 105, 275. *See*
  - also inconsistency-tolerant logics.
- paracusia, 420
- paradox-generating statements, 288
- parameter
  - mathematical definition of, 136,
    - 701
- parameters of constraint, 141-143,
  - 200, 203, 221, 325, 327-328, 336-
    - 337, 347, 356, 365-366, 396, 443,
      - 451, 509, 511-512, 517-518, 614-
        - 615, 662
    - positive and negative, 143
  - particulars, 241, 298-299, 312, 331,
    - 699-702, 704, 707, 709, 711-713,
      - 715-716, 719, 721. *See also under*
        - MoR, definition of particulars in.
    - concept of, 239-240, 701
    - definition of, 238
  - Passmore, John, 61, 110-119, 121,
    - 123, 125, 191-193, 283, 352, 359,
      - 742, 791
    - activity-based approach, 125
  - past
    - framework-independent, 445
    - “Pathologies of Everyday Thought”

- (Bartlett), 263
- pathology  
 conceptual, 6. *See also* conceptual pathology.  
 extension to normal individuals, 651
- pattern recognition, 209  
 abilities, 520, 532  
 non-algorithmic nature of, 532  
 non-formalizable, 532
- Pauli, Wolfgang, 573
- p-circularity, 178-180, 182-184, 348.  
*See also* circularity, presumptive.
- Peirce, Charles Sanders, 411
- personal identity, 494, 496-498, 501, 505, 507, 519  
 framework-relative nature of, 496  
 perspectivally given, 496
- persons. *See* personal identity. *See also* observer, role of in metalogic of reference. *See also* observer, the, concept of.
- perspectives, 230, 286  
 correlation of sensory, 425  
 definition of, 219, 224, 424  
 spatial, 448-449  
 temporal, 445  
 theory of, 286
- perturbation  
 measurement-based, 592, 594
- Petit, Jean-Pierre, 542
- petitio principii*, 206, 350
- phaenomenologia generalis*, xxxiii, 256, 279, 634, 636
- phenomenology  
 as error-detecting and error-eliminating, 634  
 experience as self-enclosed field, 335  
 first use of term, 634  
 Lambert's conception of, 634
- philosopher  
 what it takes to change the mind of, 26-32, 358
- philosophical analysis  
 as therapy, 282-283
- philosophical argument  
 formalization of  
 advantages of, 676-678  
 disadvantages of, 676, 678-680
- philosophical argumentation  
 appeal to logic, 105  
 application of self-reference, 105
- philosophical criticism  
 conceptual, 106  
 person-directed, 106  
 rhetorical, 106
- philosophical invariance. *See* invariance.
- Philosophical Investigations* (Wittgenstein), 57
- philosophical neutrality, 70, 72-73, 91, 280, 330, 514
- philosophical positions  
 as dynamic systems, 192  
 as functional organizations, 176  
 as person-based, 176  
 self-enclosure of, 335
- Philosophical Reasoning* (Passmore), 112
- philosophical systems  
 self-enclosure of, 417, 454
- philosophical universality. *See* universality.
- philosophy  
 possible future for, 668-670  
 as mother of the sciences, 1  
 as rigorous science, 8, 65. *See also* scientific philosophy.  
 disfavored views, 20  
 dispositional traits, 35  
 evolutionary stages of, 666-667  
 favored views, 20  
 ideological self-encapsulation, 30  
 internal psychological limitations of, 18, 45, 753, 755

- lack of definitive results, 7-9, 17, 24, 28, 33, 257
- lack of unified method, 105
- need for intellectual humility, 670-672
- need for method of error
  - elimination, 638
- paradigm changes in, 22
- perennial, 21
- policing function of, 22
- preferred beliefs, 28
- psychological assessment of 15-43, 671
- psychology of. *See* psychology of philosophers.
- revolutions in, 667-668
- scientific, 22. *See also* scientific philosophy.
- shifting the ground of debate, 29
- study of narratives, 21
- therapeutic, 5
- Thomistic, 119
- philosophy of language, 46-57, 59, 64-65, 88, 90, 230, 278, 284
  - of ordinary language
    - as therapy, 282
- physical identifying reference, 537, 540-541, 564, 567, 589, 592
  - in relativity physics, 544, 559
- physical laws
  - as general functional dependencies, 472
  - covariance of, 562
  - falsifiability of, 475, 564
  - invariance of, 562
  - properties of, 475
  - universal and covariant, 208
  - without reference to time variables, 472
- physical objects, 446
  - as reference structures, 99, 537, 608
  - continuity of, 435-436, 448
    - given in perspectives, 448
    - regularity of, 435-436
- physical phenomena
  - framework relativity of, 618
- physical reference frames
  - correlation of, 541
  - distinguished from coordinate systems, 528
- physics
  - epistemology of, 376
- Piaget, Jean, 582
  - concrete and formal operational thought, 37, 42
- Planck, Max, 174, 207, 573, 612
- Planck's constant, 574
- Plato, 27, 51, 287, 381-382, 657, 792, 798
  - allegory of the cave, 287
  - analogy of the line, 34
- Podolsky, Boris, 392, 394, 575
- Poincaré, Henri, 534
- Polányi, Michael, 42, 356
- polar coordinate system, 217, 220
- polemic
  - nature of, 42
- Pólya, George, 333-335, 338, 379
- Popper, Karl, 471
- positive science, 632
  - definition of, 631
- positivism, 631
  - therapeutic, 282
- possibility
  - absolute, conceptual, or *a priori*, 139
  - as complementarity, 135
  - as consistency, 134
  - as freedom from contradiction, 134. *See also* possibility, formal.
  - as self-consistency (Fitch), 134
  - concept of, 525, 707
  - concepts of, 132, 154
  - epistemic, 134, 141
  - formal, 134, 141

- framework-relative, 138-139, 141, 154, 200-201
- general theory of, 141, 280
- in terms of necessity, 143
- in terms of parameters of
  - constraint, 142, 200
- many-worlds, 137
- metalogical, 139-141
- modal concept of, 132, 154
- nomological or physical, 134, 141
- parametric, 136-137, 141
- psychologically based, 132, 141
- relative, 139
- temporally based, 132-133, 141
- theory, 131-142
- theoretically general concept of, 517
- Possibility* (Buchanan), 136
- possible
  - knowledge
    - level of, 347
  - meaning
    - level of, 347
  - worlds
    - definition of, 139
    - nature of, 138
- Post, Emil Leon, 680
- pragmatical self-referential
  - inconsistency, 242, 345-346
- pragmatics, 88
- pragmatism, 110
- preconditions
  - of identification, 142, 154, 236
  - of identification/identifying
    - reference, 225
  - of possibility, 77-78
  - of possible meaning, 250
  - of possible reference, 85, 127
  - of reference, 94, 102
  - of referring and identification, 4
  - structural/systemic, 214
    - first-order, 210
    - transcendental, 132, 146
- pre-reflective experience, 129, 493, 505
- Presley, C. F., 267
- presuppositions
  - absolute (Collingwood), 193
    - as rules of a game, 193
  - as function of identifiability, 203, 313
  - as function of meaning, 203, 313
  - as function of truth, 203
  - as ingredients of definitions, 159
  - as logically antecedent
    - suppositions, 158
  - as preliminary assumptions, 157-158
  - as suppositions of language use, 159-160
  - conditional, 158
  - destructive testing for, 320
  - linguistic of reference, 190
  - metalogical, 331
  - method to test, 172
  - necessary for possible referential
    - frameworks, 143
  - non-truth-functional, 181
  - not language-based, 172
  - of all proposing (Passmore), 112-114, 117, 191-193
  - of concepts, 163, 192, 313-314
  - of existence, 161-162, 165, 190, 313
  - of identifiability, 313
  - of identification, 163, 186, 188-190, 194, 204, 280, 312-314, 316, 328
    - as modal, 189
    - conceptually the most
      - fundamental, 186
    - framework-relative nature of, 186
    - metalogical nature of, 315
    - metalogic of, 69, 189
    - non-linguistic, 189

- structural/systemic nature of,
  - 188
  - sub-class of structural/systemic, 186
  - test for, 314, 315
- of language use, 190
- of linguistic reference, 313
- of meaning, 313
- of missing premises, 160
- of pragmatic activity, 160, 162, 190
- of reference
  - linguistic, 162, 165
- priority of, 157
- psychological, 159-160
- semantical, 159
- structural/systemic, 163, 172-176, 191, 194, 204, 210, 221-224, 253, 280-281, 301, 312-314, 316, 328, 330-331, 339, 347-349, 355, 373-374, 665
  - not truth-functional, 176
- summary of varieties of, 163-164
- tests for, 316
- truth-functional, 165, 313
- truth-functional referring, 168-171, 174
  - varieties of, 182
- Priest, Graham, 276, 400-404, 417, 458
- Principia Mathematica* (Russell & Whitehead), 681
- principle of complementarity, 584, 587, 589-590, 609, 621
- principle of equivalence, 546-547
- principle of indeterminacy, 574. *See also* uncertainty principle.
- principle of relativity, 535, 538, 542, 803
- principle of special relativity, 542
- principle of uncertainty, 205, 574, 583, 592, 610, 620
- Prior, Arthur, 260
- priority
  - logical, 4
  - transcendental, 4
- privileged introspection or access, 344, 347, 352
- probability, 209, 239, 288, 473
- problem
  - definition of, 377
  - relationship to solution, 377
- problem of induction, 471
- problem of putative meaning, 238, 247, 260-263, 275, 281, 299, 388, 694, 720
  - involving shift of significance, 262
- problem-solving
  - discovery in 372-390, 376
  - invention in 372-390, 376
- Process and Reality* (Whitehead), 76
- projection, 242, 244-252, 280, 287-288, 291, 311-312, 316, 336-347, 355-356, 359, 364-365, 374, 388, 397, 414, 428, 430, 438, 632, 657
  - allegorical origin of term, 287
  - as systemic error, 638
  - claim of framework autonomy, 298-299, 325, 328
  - concept of, 297-301
  - definition of, 298-299, 325, 327
  - epistemological component, 297
  - historical intimations of, 289-296
  - modal level of, 300
  - meaning in mathematics, 286
  - meaninglessness of, 301
  - need to reject, 329
  - negation of, 249
  - of absolute space, 449-452, 456, 560
  - of absolute time, 446, 455, 560
  - of agency, 501
  - of causal agency, 468-469, 485. *See also* projections, of genetic causation.
  - of causality, 511

of consciousness as a container, 503  
 of counterfactual causation, 466-467, 472, 480-481, 485, 500  
 of discovery, 384  
 of framework-transcending causality, 471-472, 482, 485  
 of general continuity, 435  
 of genetic causation, 468-469, 472, 481, 485, 500  
 of hidden determinants, 472, 479, 481, 485, 500  
 of idealism, 439-440, 508  
 of induction, 471  
 of invention, 386  
 of mental faculties, 499-500  
 of noumenal things-in-themselves, 430  
 of other minds, 431-434, 494, 508  
     as things-in-themselves, 433-434  
 of realism, 439, 508  
 of reflection, 493, 498, 500-501  
 of space, 456  
 of space-time, 454-456, 560  
 of spatial constitutive subjective activity, 452, 456-457  
 of spatial continuity, 601  
 of spectator consciousness, 502, 509  
 of temporal and spatial continuity, 453-454, 456  
 of temporal constitutive subjective activity, 447, 456-457  
 of temporal continuity, 601  
 of the external world, 428, 431, 450, 494, 508  
 of the future, 445-446, 455, 560  
 of the implicit, 128, 454, 479  
 of the past, 444, 455, 560  
 of the self, 511, 597  
     and role of the observer, 518  
 of the self as an existing entity, 495, 498, 500-501

of the self as bearer or owner of its states, 498, 500  
 of the self as center of experience, 497-498, 500, 561  
 of things-in-themselves, 450, 561  
 of time-flow, 560, 444-445, 455  
 of transcendence, 762-763  
 origins of term, 284-285  
 psychological component of, 297  
     in Freud, 285  
 reifying (Freud), 285-286  
 rejection of, 247-248, 250  
 relating to ontology, 374  
 rule-prescriptive (Freud), 285-286  
 severing of essential framework  
     relativity, 298-299, 325  
 symbol for, 237  
 that thinking entails a thinker, 493, 498, 500  
 varieties of, 659-660  
 variety of meanings of, 285  
 vector notation for, 237, 242, 297  
 projections  
     common in all disciplines, 655-656  
     detection of, 323  
         automation possibility, 319  
     elimination of, 323  
     method to detect, 313. *See* de-projection.  
     re-formulation of, 323  
 projective assertions. *See* projections.  
 projective belief, 650  
     psychology of, 651-654  
 projective delusion, 665, 694  
 projective geometry, 222, 286  
 projective meaninglessness. *See* projection.  
 projective misconception, 242. *See* projection.  
 projective self-referential inconsistency. *See* projection.  
     modal level of, 300  
 projective thinking

- delusional nature of, 414
- projective vector
  - magnitude and direction of, 298
- Prolegomena to Any Future Metaphysics* (Kant), 307
- propriety of statements (Palmer), 179
- Proust, Marcel, 373
- pseudoproblems (Carnap), 293
- pseudo-problems (Halldén), 270
- psychological compulsion to over-reach, 414. *See* horizons, human compulsion to transgress.
- psychological normality
  - internal limitations of, 757
- psychology
  - clinical, 32
  - faculty, 498
  - of referential thinking, 187
  - role of in this book, 5, 640-642
- psychology of normality, 26
- psychology of philosophers, 32, 645, 735-746, 749-753
  - common psychological traits, 750
  - contentiousness, 566, 744-745, 751
  - defensiveness, 30, 38, 646, 667, 671, 744-747, 751, 754, 758
    - forms of, 742
  - denial, 28, 31, 35, 44
  - distortion, 28, 29, 31, 35, 44
  - narcissism, 23, 30, 35, 81, 104, 743-744, 746-747, 754, 758
  - rejection, 28, 31, 35, 44
  - resistance to criticism, 106, 742-744, 745, 751
- Ptolemaic model, 284
- Ptolemy, Claudius, 496, 666
- putative meaning, 324, 640. *See also* meaning, putative. *See also* meaning, problem of putative.
- putative reference, 297, 327, 404, 720
- Putnam, Hillary, 60
- Pylyshyn, Z., 264
- quantum field theory, 212
- quantum logics, 248, 636
- quantum phenomena
  - autonomous reality of, 575, 582
    - projective nature of, 583
  - discontinuity, 621
  - framework relativity of, 574, 588-593, 597, 604-607, 610
  - indeterminacy of, 585, 621
  - indeterminate nature of, 584, 605
  - non-relational ontology of, 596-600
  - randomness of, 621
- quantum theory, 221, 223, 234, 284, 373, 423, 426, 428, 451, 472, 479, 523-524, 563-607
  - affinities with metalogic of reference, 611
  - as set of limitative results, 609-611, 623
  - canonically conjugate properties, 574, 576, 578, 580, 583-584, 587, 603, 606, 608, 610, 620-621, 623
  - complementarity, 577, 608
  - completeness of, 621
  - concept of objectivity, 427, 602-609, 611, 622-623
  - confirmation of results by the metalogic of reference, 615-617
  - consistent histories approach, 610
  - contributions to epistemology, 624-625
  - Copenhagen interpretation, 135. *See also* Copenhagen interpretation.
  - de-projected results of, 622-623
  - de-projective understanding of, 604-611
  - dimensionally incompatible frameworks, 608. *See also* complementarity.
  - discontinuity, 581, 600-602, 608

- epistemological basis of, 612
- hidden variable
  - interpretation, 128
  - proposals, 578-580
  - projective nature of, 579-580
  - theories, 479, 605
- incompleteness of, 576. *See also*
  - EPR (Einstein, Podolsky, Rosen) position.
- indeterminacy, 608
- inseparability of observer and
  - observed, 589-596
- measurement problem, the, 594
- measurement-based perturbation, 589, 594, 610
- mutually interfering observables, 579-580. *See also* quantum theory, canonically conjugate properties.
- ontology, 373
- perturbation. *See* quantum theory,
  - measurement-based perturbation.
  - agency-based, 590
- probability functions, 623
- projection of continuity, 602
- projection of underlying reality, 581-583
- projections involved in, 622
- projective claims of, 604-605
- randomness, 608
- realism, 575. *See also* EPR (Einstein, Podolsky, Rosen) position.
- reality given perspectively, 608
- referential fields in, 592
- renunciation of classical concepts, 588, 620-621
- renunciation of traditional
  - causality, 588
- S-matrix [scattering matrix]
  - interpretation, 598
- understanding of quantum reality,
  - 589, 608
  - variant interpretations of, 573
- questions
  - meaningful
    - necessary conditions of, 411
  - reasons they may be unanswerable, 410
  - unanswerable, 419
  - unanswerable in principle, 410. *See also* erotetic intractability.
- Quine, W. V. O., 100, 126, 209, 215-216, 349, 699, 718
  - argument with a form of closed curve in space, 210, 349
  - coordinative nature of reference, 100
  - reference requires coordinate system, 140
  - taste for desert landscapes, 49, 55-56, 60
- Ramsey, Frank Plumpton, 242
- rational acceptability
  - rules of, 357, 434
- rational admissibility
  - rules of, 359, 365, 655
- rational bridge, 645, 652, 760
- rational bridge problem, 25, 28, 357, 359, 645, 652, 760
- rational coherence, 347
- rationality, 123
  - and conviction, 26, 357-358, 645, 652, 760, 763
  - conceptual, 353, 357
  - criteria of, 354
  - epistemological, 356, 358.
  - necessary conditions of, 355
  - of the agent, 357
  - presupposition of, 346
  - psychological, 353, 357
  - rules of admissibility 351-359
  - strongly compelling criterion of, 355



- reactive boundaries, 336. *See under* horizons.
- readiness potential, 484
- realism, 128, 416, 436-439, 489, 666  
 argument against (Royce), 291-292  
 event-realism, 436-438  
 in mathematics, 382, 386  
 meaninglessness of (Carnap), 293  
 metalogically self-undermining.  
*See* projection, of realism  
 projection of, 439. *See also* projection, of realism.  
 projective nature of, 576
- reality  
 autonomous  
 belief in, 646  
 epistemologically inaccessible, 429  
 framework-relative understanding of, 224  
 macrophysical  
 discontinuous nature of, 600
- reasoning  
 human  
 simulations of, 691  
 metamathematical, 684, 686, 688  
 temporal structure of, 688  
 syllogistic model of, 182
- recalcitrance, 30-31, 35, 81, 104, 414, 646-647, 653, 663, 758  
 of system of belief, 30  
 psychological, 42  
 reflex-arc, 30, 39
- recognition  
 of commonality, 740  
 inescapable ambiguity of, 531, 739-740
- Recordé, Robert, 303
- recursion, 687-688
- recursive definition, 686
- recurvature, 61, 80-81, 195, 199, 321, 337, 349, 419, 454. *See* Moebius strip. *See also* self-enclosure.
- redargutio elenchica* (St. Thomas Aquinas), 120.
- reductio ad absurdum*, 333-335, 401
- reduplicative paramnesia, 64
- reference. *See also* identification. *See also* ostensive reference.  
 20th century view of, 59  
 activity-based, 92, 102, 187  
 agentless, 512-514, 521-522  
 as a four-place relation (Strawson), 67  
 as conceptually basic, 69  
 as coordination, 101  
 as identification, 96, 101  
 as identity-recognition, 94  
 as identity-specification, 94  
 as most basic concept, 77  
 as ternary relation, 240  
 autonymous, 13  
 boundaries of possible, 247. *See* horizons.  
 causal notion of, 72  
 concept of  
 freed from subject-object conception, 95, 511-512  
 conventional notions of  
 result in projection, 513  
 coordinative nature of, 213  
 delusional, 81, 413, 661  
 delusions of, 63  
 de-projected concept of, 514  
 indirect, 13  
 intrinsic, 548-551, 559, 563, 619  
 de-projective concept of, 551  
 iterable, 442  
 iterative, 93, 716  
 levels of, 93, 214  
 linguistic, 60, 64, 87-88, 90, 94-95.  
*See also* philosophy of language.  
 logical priority of, 277  
 meta-level, 214  
 metalogical concept of, 88, 102  
 metalogical preconditions of, 242, 300

metalogically self-undermining,  
 142. *See* projection.  
 metatheory of, 206. *See under*  
 metalogic of reference.  
 minimalist conception of, 96  
 naive conception of, 240  
 naive view of, 89-91, 94, 101  
 negative and positive constraints  
 of, 154  
 non-linguistic, 62-63, 95  
 non-relational, 87, 95, 280, 512-  
 514, 521-522, 597  
 agentless, 489  
 objects of, 91-93, 96, 101-102, 140,  
 159, 163, 186, 200, 202, 209,  
 213-217, 219, 221-223, 281,  
 298-299, 312, 315, 319, 372,  
 374-375, 398-399, 419, 442-443,  
 513, 520-521, 526-528, 591  
 ostensive, 63, 87-88, 92, 277  
 philosophy of language, approach  
 to, 65  
 positive and negative constraints  
 of, 190  
 framework-relative  
 understanding of, 200, 522  
 pragmatic, 88, 92  
 preconditions of, 330. *See also*  
*under* preconditions.  
 basis for identifiability, 200  
 property-based, 415  
 psychological approach to, 65  
 putative, 297, 327, 404, 720  
 recursive character of, 686, 688  
 reflective  
 temporal preconditions of, 492-  
 493  
 re-identifying, 260, 442, 494-495,  
 519, 531  
 reiterated, 209, 442  
 relation to identity, 101  
 relational conception of, 87  
 retrospective second-order, 209,

442  
 rule-based, 415  
 separation of from object of  
 reference, 91  
 stripped of pragmatical  
 associations, 95  
 temporal basis of, 443, 448, 455,  
 463, 686  
 temporally retentive, 450  
 temporally successive, 209, 442,  
 450. *See also* reference,  
 reiterated  
 to future events, 445  
 to past events, 444  
 to physical objects, 450  
 to spatial objects  
 temporally discontinuous, 450  
 varieties of, 60, 63-64, 67  
 reference frames  
 compatible, 218, 424, 526, 531-532  
 correlation of, 531  
 incompatible, 526  
 inertial, 207, 538, 542-547, 550,  
 553, 557-560, 564, 618  
 in physics, 208, 527  
 pre-analytical, 254  
 presupposed by analysis, 255  
 translations between, 218  
 reference mollusk. *See* Einstein's  
 mollusks.  
 referential closure, 338  
 referential consistency, 85  
 as a criterion of meaning, 79, 130,  
 312, 203, 231-232, 236-237,  
 248, 252, 278-279, 281, 296,  
 333, 347, 355, 518, 626, 763  
 intrinsically determined, 235  
 non-arbitrary nature of, 232  
 self-validating, 252  
 strongly compelling, 252  
 strongly non-arbitrary, 252  
 unique properties of, 232  
 referential constraint

- parameters of, 203
- referential fields, 213-216, 224, 312, 319, 336, 418-419, 510, 514, 516-518, 521-522, 546, 563, 591, 597, 598, 638, 664-665, 688
- as field of possibilities, 522
- concept of, 514, 515
- finitary limits of, 312
- gravitational, 548
- instantiation of, 214, 312, 319, 418-419, 516, 591
- metalogical horizon of, 508. *See also under* horizons.
- modal nature of, 522
- ontological implications of, 215
- properties of, 591
- reactive boundaries of, 312, 591. *See also under* horizons.
- relativistically self-enclosed, 522
- self-enclosure of, 216, 418, 508, 591, 683. *See also* self-enclosure.
- reflection
  - perturbation question, 129
  - temporal nature of, 442
  - temporal preconditions of, 493
- reflexive analysis, 77
- reflexive proof, 77
- reflexive theory, 77, 81, 217, 335, 347, 418, 454, 508, 522, 674
- reflexivity, 61, 103, 104, 349. *See also* self-reference.
  - biological, 61
  - in anthropology, 61
  - in artificial intelligence, 61
  - in game and decision theory, 61
  - in law, 61
  - in neuropsychiatry, 62
  - in physical theory, 61
  - in topology, 61
  - metalogical, 127. *See also* metalogical self-reference.
  - set-theoretical, 61
  - sociological, 61
  - total, 685, 689, 695
    - possible inescapable limitation of, 694
- refutation
  - pragmatic (Mackie), 116. *See also* operational self-refutation.
- Rehg, William R., 202
- Reichenbach, Hans, 22, 97-101, 194, 211, 214, 266, 337, 352, 537, 608, 672, 794
- re-identifiability, 443
  - of objects, 209
- re-identification, 442, 686
- rejection
  - metalogical, 300
- relation
  - meaning of, 597
- relativism, 17, 20, 205, 208
  - arguments against, 206
  - conceptual, 6
  - cultural, 6, 205
  - epistemological, 205
- Relativistic Theory of Phenomenological Constitution, A*, (Bartlett), 100
- relativity
  - functional, 443
  - opposed to relativism, 535
  - principle of. *See* principle of relativity.
- relativity physics, 284, 521-555. *See also* special theory of relativity. *See also* general theory of relativity.
  - affinity with metalogic of reference, 536
  - central role of covariance, 557-559
  - central role of invariance, 556-559
  - concerned with absolute invariants, 535
  - confirmation of results by the metalogic of reference, 615-617
  - contrast with metalogic of

- reference, 563
- contributions to epistemology, 624-625
- epistemological basis of, 612
- epistemological in nature, 533, 536, 617
- framework-relativity of, 99
- heuristic nature of, 563
- inappropriate naming of, 207, 534-536, 583
- invariants of, 207
- meaninglessness of pre-relativity notions, 565
- methodological principles of, 562
- naming of, 207. *See also* under *Invariantentheorie*, 207. *See also* relativity physics, inappropriate naming of.
- rejection of Newtonian tenets, 564
- rejection of operationalism, 540, 565
- renunciation of classical concepts, 618-619
- results reached by, 560
- role of coordinate transformation in, 537
- relevance logics, 636
- religion, 30, 40, 230, 504, 641. *See also* under horizons, human compulsion to transgress. *See also* beliefs, privileging of.
- Remembrance of Things Past* (Proust), 373
- Rescher, Nicholas, 410-411, 413-414
- respectful sympathy, 323-324, 359, 373, 485
- rétorsion (Isaye), 119-125, 130, 783
  - and Kantian transcendental preconditions, 125
  - not *ad hominem*, 121, 123
  - not pragmatically self-referential, 124
- retortion. *See* rétorsion.
- retroactive application of analysis, 255
- revolutions
  - conceptual, 284, 666-668
- Ricci-Curbastro, Gregorio, 547, 552
- Richman, R. J., 235
- Ricoeur, Paul, xxix, xxxv, 851
- Riemann curvature tensor, 555
- Riemann, Georg Friedrich Bernhard, 547, 552
- Robinson, Richard, 69
- Romanes Principle, 756-757, 760
- Romanes, George, 291, 756-757
- Rosen, Nathan, 392, 394, 575
- Rosenfeld, Léon, 573
- Rosser, J. Barkley, 680, 685
- Routley, Richard, 62, 272-273, 275-276, 279, 374
  - matrices for connectives, 274
- Royce, Josiah, 291-294, 296, 348, 438
- Rucker, Derek, 38
- rules
  - injunctive, 200, 522
  - prescriptive, 200, 522
- Russell, Bertrand, 58, 60, 227, 371, 404, 422-423, 427, 433, 452, 470, 473, 491, 534-535, 551-552, 655-656, 667-668, 672-674, 681, 776, 780, 783-784, 796-797
- Russell's paradox, 61, 721
- Ryle, Gilbert, 257, 282
- Saari, Heikki, 193
- Sagnac Effect, 542
- sameness (judging commonality of objects), 221, 424-425, 442-443, 448, 457, 526, 528-532, 541, 606, 686. *See* recognition, of commonality. *See also* theorem of the ugly duckling. *See also* similarity relations, inescapable ambiguity of.
- Sapir, Edward, 52

- scalar field, 553, 555  
 scalars, 553-554  
 Schiffer, Stephen, 90  
 Schilpp, Paul Arthur, 23, 33  
     that philosophers do not want to  
     understand one another, 23  
 Schlesinger, George N., 679  
 Schlick, Moritz, 234, 266, 268, 468,  
     491, 505, 672  
     murder of by student, 468  
 Schmidt, Paul F., 235  
 Schopenhauer, Arthur, 430  
 Schrödinger, Erwin, 573, 593  
 Schrödinger wave equation, 592  
     collapse of, 593-596  
     description of, 593-594  
 Schulze, Gottlob Ernst, 430  
 Schweitzer, Albert, 670  
 science  
     definition of, 632  
     norms of, 669-672  
 science of science (Husserl), 76  
 scientific philosophy, 420, 669-672,  
     675  
 self, the  
     activities of, 500  
     as ego, 495  
     as bearer of its states, 497-498  
     as causal agent of thinking, 499-  
     501, 509  
     as center of experience, 496-497,  
     505  
     as existing entity, 494-495  
     as framework-autonomous entity,  
     501  
     as superego, 495  
     as universal characteristic of  
     experience, 501-502  
     commonsense notions of, 490, 500,  
     507, 518  
     de-projective understanding of,  
     507-510  
     disconfirmation of, 31  
     phenomenological approach to, 491  
     possessor of faculties, 498-499  
     problem of, 489  
     subjectless experience, 492-493,  
     502, 505-507  
     that thinking entails a thinker, 491-  
     492  
     projection of, 493  
     transcendental, 495  
     transcendental unity of  
     apperception, 495  
 self-defeating concepts, 6  
 self-enclosure, 80, 81, 213, 216-217,  
     312, 335-337, 349, 420, 454, 508-  
     509, 591, 664  
     concept of, 417-419, 522  
     in restricted sense, 217  
     of delusional belief  
         in psychiatric patients, 418  
     of experience  
         in phenomenology, 81  
     of maximally general theory, 211  
     of metalogic of reference, 217  
     of philosophical systems, 216  
     of systems of belief, 216  
     relativistic, 418  
     systemic, 81  
     topological, 81. *See also* Möbius  
     strip.  
 self-evidence, 343, 352  
 self-reference, 61, 103. *See also*  
     reflexivity.  
     in academic scholarship, 62  
     in art, 61  
     in economics, 61  
     in fractal theory, 61  
     in hermeneutics, 61  
     in information theory and general  
     systems theory, 61  
     in literature, 61  
     in music, 61  
     in mythology and theology, 61  
     in political science, 61

- in psychiatry and psychology, 61
- in science fiction, 61
- in theory of paradigm change, 61
- in theory of research programs, 61
- language-based
  - varieties of, 103
- metalogical, 242
  - constructive use of, 129-130
  - examples of, 127
- of statements vs. propositions, 104
- performative, 103
- pragmatical (or performative), 61, 103, 129, 241, 481
  - uses of, 110
- propositional, 104
- semantical, 61, 103
- sentential, 241
- tautological, 61
- varieties of, 61
- self-referential argumentation, 103-104. *See also* argumentation, self-referential.
  - absolute (Passmore), 125. *See also under* Passmore, John.
  - absolute vs. pragmatical, 114
  - performative or pragmatical, 108
- self-referential inconsistency. *See specific kinds.*
  - metalogical, 126, 141, 201. *See also under* projection.
  - performative, 345
  - pragmatical, 346
  - rétorsion
- self-referential index, 840
- self-referential theory, 75-76, 663, 675
- self-refutation
  - absolute
    - in Mackie, 116
    - in Passmore, 112-114
  - ad hominem* (Passmore), 112
  - as tool only of negative criticism (Mackie), 118
  - operational (Mackie), 115-116, 118-119
  - varieties of, 116
  - philosophically significant use of, 117
  - pragmatic (Mackie), 116-117
  - pragmatical (Passmore), 112-113
- self-sustainment, 344-348, 352
- self-validating proofs, 77. *See also* self-validation.
- self-validation, 129, 201, 247-250, 323, 330, 339-350, 374, 657, 706.
  - irrefutability of, 350-351
  - logic of, 347-350
    - recurvature of, 349
- self-verification, 344-347, 352
- semantic theory, 103
- semantically closed language, 338
- semantics, 54, 88, 144
  - of natural language, 228
- semiotical psychiatry, 282
- semiotical therapy, 282
- sensory experience, 432-435, 440, 448, 470, 634
  - continuity and regularity of, 453
  - discontinuity of, 601
- sentences (or propositions)
  - putatively referring, 238
  - putatively true or false, 243
- Sesonske, Alexander, 26
- Sextus Empiricus, 401
- Shaw, George Bernard, 317
- short-circuit
  - metalogical, 142, 203, 218, 246, 265, 313, 325, 327, 329, 372, 459, 466, 651, 688
  - truth-functional, 126
- signal processing, 223
- significance logic, 273
- similarity relations. *See* Theorem of the Ugly Duckling.
  - inescapable ambiguity of, 464, 531

- Simon, H., 317
- simultaneity, 560  
 absolute, 560  
 relativity of, 542
- skepticism, 110, 163, 202, 206  
 arguments against, 196  
 pragmatical self-referential  
 arguments against, 196  
 Strawson's position against, 128
- Skinner, B. F., 379
- Skolem, Thoralf Albert, 680
- Smith, Norman Kemp, 307
- Social Security numbers as identifiers,  
 60, 186, 702
- sociology, 355
- Socrates  
 role of conviction, 26
- solipsism, 489-490, 502-505, 507,  
 509, 801  
 problem of, 489  
 projections involved in, 505  
 without a self, 509-510
- Salovey, Peter, 737
- sophistry, 107, 359
- space  
 as continuum, 449  
 concept of 438-454  
 de-projective understanding of,  
 457-458
- space-time  
 concept of 438-454  
 curvature of  
 functionally defined, 476  
 relativistic conception of, 455
- spatial objects, 443  
 identifiability of, 448  
 identifiability temporally based,  
 455
- spatial order, 447-449  
 belief it is everywhere dense, 453-  
 454, 478  
 presupposes temporal, 448-449  
 serial nature of, 448
- special theory of relativity, 208, 218-  
 219, 472, 523, 528, 532-533, 537-  
 539, 542-543, 545, 556, 564, 566  
 definition of length, 541, 543  
 definition of time, 541, 543  
 invariance of physical laws, 557  
 length contraction, 546  
 measurements of distances and  
 times, 544  
 measurements of durations, 544  
 reference-based consequences of,  
 544  
 specification of simultaneity, 544  
 time dilation, 546
- species problem, 209, 464, 532, 702,  
 739
- spectator theory, the, 501
- spherical coordinate system, 217
- Spinoza, Baruch, 534
- square roots  
 manual extraction of, 191
- St. Augustine (Augustine of Hippo),  
 51
- Stapp, Henry, 597
- Stebbing, Lizzie Susan, 87
- Strawson, P. F., xxxv, 60, 67, 90, 128,  
 130, 165, 194, 202, 497, 700, 783,  
 800
- string theory, 626
- Stroll, Avrum, 259
- Strong-Campbell Interest Inventory,  
 30
- Stroud, Barry, 202
- subatomic particles  
 as excitations of fields, 516
- Suber, Peter, 25
- subjectivism, 110  
 radical. *See also* solipsism.
- subject-object. xxxiv, 511, 513, 593,  
 597, 664, 667. *See also* reference,  
 concept of, freed from subject-  
 object conception.
- successor function, 686

- superposition  
 in classical wave mechanics, 594  
 in quantum theory, 594
- Suppes, Patrick, 249, 272
- Sur la calculabilité du nombre de toutes les connaissances possibles* (Leibniz), 304
- syllogism, 338-339  
 logical structure of  
 autarkic, 338-339. *See also* autarky.
- sympathy  
 respectful, 323, 329, 373, 485  
 meaning of, 324
- synchronous coordinates. *See* Gaussian coordinates.
- syntax, 88
- systems of reference  
 field-based nature of, 518. *See also* under referential fields.
- Tagiuri, R., 737
- Tarski, Alfred, 78, 680, 685  
 definition of truth, 722
- tautology, 199  
 in logic, 330  
 in mathematics, 330  
 meaning of, 336  
 word origin, 330
- taxonomy  
 Linnaeus, Carl, 70  
 psychological predispositions in, 746-747
- Taylor, C. W., 379
- Teller, Edward, xxxv
- temporal order, 448  
 serial nature of, 448
- temporal succession, 448  
 belief it is everywhere dense, 453-454, 478  
 of events, 445
- Ten Commandments, 143, 718
- tensors  
 concept of, 551-556  
 Einstein, 555  
 etymology of the word, 552  
 Ricci, 555  
 Riemann, 555
- tensor field, 555  
 concept of, 551-556
- Terrell, G., 379
- Theorem of the Ugly Duckling (Watanabe), 740
- Theorema Egregium, 548, 551
- theorems of formal limitation, 684-685, 688-690  
 described recursively, 687  
 reliance upon metamathematical reasoning, 688, 693. *See also* metamathematical reasoning.
- theories  
 as systems of coordination, 443
- theory  
 of argument, 103. *See specific kinds of argument.*  
 of classes, 103  
 of invariance. *See Invariantentheorie.*  
 theory of knowledge, 97, 103, 534, 776. *See also* epistemology.
- theory of theories (Husserl), 76
- theory of types, 249, 721
- therapy  
 conceptual, 5. *See* conceptual therapy.  
 linguistic, 1  
 philosophical, 5
- things-in-themselves, 429. *See* projection, of things-in-themselves.
- Thomism  
 transcendental, 731
- Thorndike, E. L., 737
- thought  
 beyond limits of thought, 400-402  
 projection of, 401
- thought disorders, 648-651



- derailment, 650
- extending to normality, 650, 654
- illogicality, 650
- three-valued logic, 249, 288
  - without a theory of types, 249
- time
  - concept of 438-454
  - de-projective understanding of, 456-458
  - passage of, 444
- time dilation, 542
- time-order, 444, 447, 540, 544, 688
- Todd, William, 643
- token-reflexives, 61
- Tollefsen, Olaf, 104, 481
- Tolstoy, Lev Nikolayevich (Leo), 668
- topology, 195, 199, 216, 321, 337, 419
  - of space-time manifold, 454
  - recurvature, 321. *See* recurvature.
    - See also* Moebius strip. *See also* de-projection, method of, recurvature of. *See also* self-validation, recurvature of. *See also* Worm of Ouroborus.
  - referential, 420
- Toulmin, Stephen, 282
- Tractatus Logico-philosophicus* (Wittgenstein), 57
- transcendence
  - belief in, 762
  - a disorder of thought, 762
- transcendent objects, 197
- transcendental analysis, 70
- transcendental arguments, xxxiii. *See* transcendental argumentation. *See also* argument, transcendental.
  - objective of, 131
  - presumptive circularity of, 178. *See also* circularity, presumptive.
- transcendental argumentation, 131-132, 145-146, 176, 178, 184, 194-202, 293, 300, 337, 348-349, 419, 522, 633, 731
  - against denial of other minds, 196
  - and syllogistic logic, 197
  - as self-validating, 199
  - as structural/systemic, 198
  - as tautological, 101, 199
  - by means of metalogical self-reference, 196
  - central objective of, 200-201
  - circularity of, 184, 195, 197, 210, 337, 348, 349, 419
  - dynamic character of, 198
  - main approaches to, 195
  - purpose of, 194
  - recurvature of, 195. *See also* self-validation, recurvature of.
  - shortcomings of, 197
  - steps of, 197
  - sylogistically expressed, 198
    - failure of, 198
  - systems-based, 349
  - toward a general theory of meaning, 202
- transcendental deduction, 85
  - in Kant, 130, 196
- transcendental horizon, 309. *See also* *under* horizons.
- transcendental phenomenology (Husserl), 130
- transcendental philosophy, 85
- transcendental preconditions
  - of objective knowledge, 104
- transcendental proof (Kant), 176-178
- transcendental self-referential argumentation, 280
- transformation of coordinates, 217-221, 424
- truth
  - coherence theory of, 110
  - incontrovertible, 344
  - pursuit of, 341
  - Tarski's definition of, 244
- truths

- unknowable, 419
- unknown. *See* Fitch's Theorem.
- twin prime conjecture, 383
- types
  - theory of, 249, 721
- uncertainty principle, 205, 574, 583, 592, 610, 620. *See also*
  - indeterminacy. *See also* quantum phenomena, indeterminacy of.
  - inappropriate naming of, 583
- unintelligibility, 356
- universality, 20, 68, 70, 72-73, 82, 207, 280, 534, 564, 603, 781
- unprovability and incompleteness philosophical, 45
- use-mention distinction, 240
- utilitarianism
  - arguments defending, 112
  - views that oppose, 110
- van Fraassen, B., xxxiii
- Vasil'ev, 135, 774, 784, 802
- vector, 554
  - meaning in mathematics and physics, 297
  - phenomenological concept of, 298
  - projective, 297. *See* projection, vector notation for.
  - direction of, 297
- vector field, 554
- verifiability
  - arguments defending as a criterion of meaning, 257
- verital theories, 75-77, 81, 217, 335, 347, 418, 508, 603, 663-675
- Vienna Circle, 257
- von Humboldt, Wilhelm, 52
- von Neumann, John, 579
- von Schubert-Soldern, Richard, 491
- von Weizsäcker, Carl Friedrich, xxix-xxxi, xxxiii, xxxv-xxxvi, 571-572, 582, 589, 597, 599-600, 612, 623, 697, 802
- von Wright, G. H., 148
- Wang, Hao, 678, 680, 685
- Ward, Lester Frank, 291
- Watanabe, Satosi, 740
- Weiler, Gershon, 257
- Weinberg, Steven, 212-213, 534
- Wertheimer, M., 379
- Whitehead, Alfred North, 76, 214, 449, 681
- Whorf, Benjamin Lee, 52
- willful blindness, 38-39, 43, 489, 582, 639, 663, 781
- Will to Believe* (James), 36
- Williamson, Timothy, 8, 42, 50, 54, 90
- Winternitz, Josef, 534
- Wisdom, John, 282-283
- Wittgenstein, Ludwig, xxxiii, 5, 57-59, 102, 135, 211, 214, 227, 233, 282-283, 403, 417, 419, 505, 509, 699, 727, 797, 801, 803-804
- Wolpert, Lewis, 30
- Woods, J., 275
- worldview
  - strongly compelling, 420
- Worm of Ouroboros, 346, 349, 351
- Youngren, Mary Ann, xxxvii
- Zaner, Richard M., 74-75
- Zermelo, Ernst, 337
- Zermelo-König paradox, 61
- Zschimmer, E., 207
- Zucker, Hans, xxix, xxxvi

## About the Author

Homepage

Steven James Bartlett was born in Mexico City and educated in Mexico, the United States, and France. His undergraduate study was at the University of Santa Clara and at Raymond College, an Oxford-style honors college of the University of the Pacific. He received his master's degree from the University of California, Santa Barbara; his doctorate from the Université de Paris, where his research was directed by Paul Ricoeur; and he has done post-doctoral study in clinical psychology and psychotherapy at Saint Louis University and Washington University. He has been the recipient of many honors, awards, grants, scholarships, and fellowships. His research has been supported under contract or grant by the Alliance Française, the American Association for the Advancement of Science, the Center for the Study of Democratic Institutions, the Lilly Endowment, the Max-Planck-Gesellschaft, the National Science Foundation, the Rand Corporation, and others.

Bartlett brings to the present work an unusual background consisting of training, research, and publication in epistemology, philosophy of science, mathematical logic, and clinical psychology. He is the author or editor of more than twenty books and monographs, and many papers and research studies in the disciplines of epistemology, philosophy of science, mathematical logic, and psychology. He has held professorships at Saint Louis University and the University of Florida, and research positions at the Max-Planck-Institute in Starnberg, Germany and at the Center for the Study of Democratic Institutions in Santa Barbara. He holds honorary faculty research appointments from Willamette University and Oregon State University, and devotes full-time to research and publication.





