The operationalization of general hypotheses versus the discovery of empirical laws in Psychology

Stéphane Vautier

The operationalization of general hypotheses versus the discovery of empirical laws in Psychology

Stéphane Vautier
Université de Toulouse

Abstract

Psychology students learn to operationalize 'general hypotheses' as a paradigm of scientific Psychology: relatively vague ideas result in an attempt to reject the null hypothesis in favour of an alternative hypothesis, a so-called research hypothesis, which *operationalizes* the general idea. Such a practice turns out to be particularly at odds with the discovery of empirical laws. An empirical law is defined as a nomothetic gap emerging from a reference system of the form $\Omega \times M(\mathbf{X}) \times M(\mathbf{Y})$, where Ω is a set of events or dated objects for which some states in the set $M(\mathbf{Y})$ are hypothetically impossible given some initial conditions depicted in the set $M(\mathbf{X})$. This approach allows the knowledge historian to carefully scrutinize descriptive and nomothetic advances in contemporary empirical Psychology.

Résumé

L'enseignement de la méthodologie scientifique en Psychologie confère un rôle paradigmatique à l'opérationnalisation des "hypothèses générales" : une idée sans rapport précis à l'observation concrète se traduit par la tentative de rejeter une hypothèse statistique nulle au profit d'une hypothèse alternative, dite de recherche, qui *opérationnalise* l'idée générale. Cette démarche s'avère particulièrement inadaptée à la découverte de lois empiriques. Une loi empirique est définie comme un trou nomothétique émergeant d'un référentiel de la forme $\Omega \times M(\mathbf{X}) \times M(\mathbf{Y})$, où Ω est un ensemble d'événements ou d'objets datés dont certains états dans l'ensemble $M(\mathbf{Y})$ sont par hypothèse impossibles étant données certaines conditions initiales décrites dans l'ensemble $M(\mathbf{X})$. Cette approche permet de préciser le regard que l'historien des connaissances peut porter sur les avancées descriptives et nomothétiques de la Psychologie empirique contemporaine.

This article is the result of the author's need to elaborate on the persistent dissatisfaction he feels with the methodology of scientific research in Psychology, and more precisely with his perception of the way in which it is taught. It would indeed be presumptuous to present the following criticism as being a criticism of *the* methodology of scientific research in Psychology as a whole, since the latter is a notion which is too all-encompassing in its scope to serve as a precise description of the diversity of research practice in this vast field. The source of this dissatisfaction is to be found in what Reuchlin (1992) calls the 'distance' between 'general theory' and a 'specific, falsifiable hypothesis' (p. 32). A certain form of academism shapes the approach to scientific research in Psychology according to a three-stage process for the formulation of hypotheses (e.g., Charbonneau, 1988). When they write the report of an empirical study, researchers in Psychology must supply the grounds for their research by introducing a so-called general (or theoretical) hypothesis, then show how they have tested this hypothesis by restating it as a so-called operational (or research) hypothesis. In principle, this restatement should involve data analysis, finalised by testing at least one inferential statistical hypothesis, the so-called null hypothesis.

As a socially regulated procedure, the sequencing of theoretical, operational and null hypotheses – which we refer to here as *operationalization* – may not pose scientific problems to researchers who are mainly concerned with adhering to a socio-technical norm. The sense of dissatisfaction arises when this desire for socio-technical compliance is considered in the light of the hope (albeit an admittedly pretentious or naïve hope) of discovering one or more empirical laws, i.e. demonstrating at least one, corroborated general empirical statement (Vautier, 2011).

With respect to the discovery of empirical laws, operationalization may be characterised as a paradigm, based on a 'sandwich' system, whose workings prove to be strikingly ineffective. The 'general hypothesis' (the uppermost layer of the 'sandwich' system) is not the statement of an empirical law, but a pre-referential statement, i.e. a statement whose empirical significance has not (yet) been determined. The null hypothesis test (the lower layer of the 'sandwich') binds the research procedure to a narrow, pragmatic decision-making approach amid uncertainty – rejection or acceptance of the null hypothesis – which is not germane to the search for empirical laws if the null hypothesis is not a general statement in the strict sense of the term, i.e. held to be true for all the elements in a given set. Between the external layers of the 'sandwich' system lies the psychotechnical and statistical core of the operationalization paradigm, i.e. the production of psychological measurements to which the variables required for the formulation of the operational hypothesis are linked. Again, the claim here is not that this characterisation of research procedure in Psychology applies absolutely universally; however, operationalization as outlined above does appear to be sufficiently typical of a certain orthodoxy to warrant a thorough critical analysis.

This paradigm governs an approach which is destined to establish a favourable view of 'general hypotheses' inasmuch as they have psychotechnical and inferential support. However, the ideological interest of these statements does not automatically confer them with nomothetic import. Consequently, one cannot help wondering whether the rule of operationalization does not in fact serve to prevent those who practise it from ever discerning a possible historical failure of orthodox Psychology to discover its own empirical laws, by training the honest researcher not to hope for the impossible. After all, we are unlikely to

I wish to express my thanks to Nadine Matton and Éric Raufaste for their helpful comments on a previous version of this article. This work was funded in part by the ANR-07-JCJC-0065-01 programme. Correspondence about this article should be sent to Stéphane Vautier, Université de Toulouse, OCTOGONE-CERPP, Pavillon de la Recherche, 5 allées A. Machado, 31058 Toulouse Cedex 9, France. E-mail: vautier@univ-tlse2.fr.

worry about failing to obtain something which we were not looking for in the first place. We shall see that an empirical law consists precisely of stating an empirical impossibility, i.e. a partially deterministic falsifiable statement. As a result, we have inevitably come to question psychological thought as regards the reasons and consequences of an apodictic approach to probabilistic treatment of the empirical phenomena which it is investigating.

This article comprises four major parts. First of all, we shall illustrate operationalization on the basis of an example put forward by Fernandez and Catteeuw (2001). Next, we shall identify two logical and empirical difficulties which arise from this paradigm and demonstrate that they render it unsuitable for the discovery of empirical laws, then detail the logical structure of these laws. Lastly, we shall identify some methodological guidelines which are compatible with an inductive search for partial determinisms.

An example of operationalization: smoking cessation and anxiety

Fernandez and Catteeuw (2001) put forward the following sequence (p. 125):

> General hypothesis: undergoing smoking cessation tends to increase anxiety in smokers rather than reduce it.
>
> ↓
>
> Operational hypothesis: smokers undergoing smoking cessation are more prone to anxiety than non-cessation smokers.
>
> ↓
>
> Null hypothesis: there is no difference between anxiety scores for smokers undergoing smoking cessation and non-cessation smokers.

This example can be expanded so as to offer more opportunities to engage with the critical exercise. There is no difficulty in taking Fernandez and Catteeuw (2001) operational hypothesis as a 'general hypothesis'. Their formulation specifies neither the empirical (nominal) meaning of the notion of smoking cessation, nor the empirical (ordinal or quantitative) significance of the notion of anxiety, even though it makes reference to the ordinal operator *more prone to anxiety than*; lastly, the noun *smokers* signifies only an indefinite number of people who smoke.

The researcher may have given herself/himself a set of criteria which is sufficient to decide whether, at the moment when they examines an individual, the person is a smoker or not, and if they is a smoker, another set of criteria sufficient to decide whether or not they is undergoing smoking cessation. These sets of criteria allow the values for two nominal variables to be defined, the first attributing the value of *smoker* or *non-smoker*, and the second, which is conditional on the status of 'smoker', attributing the value of *undergoing cessation* or *non-cessation*. However, the statistical definition of the 'undergoing cessation' variable requires a domain, i.e. elements assigned a value according to its codomain, the (descriptive) reference system of the variable: {undergoing cessation, non-cessation}. The researcher may circumscribe the domain to pairs (smoker, examination date) which they already has obtain or will obtain during the course of their study, and thus define a so-called independent nominal variable.

They then needs to specify the function which assigns an anxiety score for each (smoker, examination date) pair, in order to define the 'anxiety score' statistical variable, taken as the dependent variable. The usual solution for specifying such a function consists in using the answers to an anxiety questionnaire to determine this score, according to a numerical coding rule for the responses to the items on the questionnaire. Such procedures, in which standardised observation of a verbal behaviour is associated with the numerical coding of responses, constitute one of the fundamental contributions of psychotechnics (or psychological testing) to Psychology; it enables anxiety means conditional on the values of the independent variable to be calculated, whence the operational hypothesis: smokers undergoing smoking cessation are more anxious than non-cessation smokers.

The operational hypothesis constitutes a descriptive proposition whose validity can easily be examined. However, to the extent that she/he considers their sample of observations to be a means of testing a *general* hypothesis, the researcher must also demonstrate that the mean difference observed is *significant*, i.e. rejects the null hypothesis of the equality of the means for the statistical populations composed of the two types of smokers, using a probabilistic procedure selected from the available range of inferential techniques, for instance Student's t-test for independent samples. Only then can the operational hypothesis, considered in the light of the two statistical populations, acquire the status of an alternative hypothesis with respect to the null hypothesis.

Now, let us restate the sequence of hypotheses put forward by Fernandez and Catteeuw (2001) thus:

> General hypothesis: smokers undergoing smoking cessation are more anxious than non-cessation smokers
>
> ↓
>
> Operational hypothesis: given a pair of variables ('undergoing cessation', 'anxiety score'), mean anxiety conditional on the *undergoing cessation* value is greater than mean anxiety conditional on the *non-cessation* value.
>
> ↓
>
> Null hypothesis: the two conditional means are equal.

Operationalization criticised

The example which we have just developed is typical of operationalization in Psychology, irrespective of the experimental or correlational nature (Cronbach, 1957, 1975) of the study. In this section, we make two assertions by dealing with the operationalization approach in reverse: (i) the empirical relevance of the test of the null hypothesis is indeterminate (ii) the statistical fact of a mean difference has no general empirical import.

The myth of the statistical population

To simplify the discussion, let us suppose that the researcher tests the null hypothesis of the equality of two means using Student's t procedure. The issue at stake in the test from a socio-technical point of view is that by qualifying the difference observed as a *significant* *difference*, the cherished notation "$p < .05$" or "$p < .01$" may be included in a research paper. The null hypothesis test has been the subject of purely statistical criticisms (e.g., Krueger, 2001; Nickerson, 2000) and it is not within the scope of this paper to draw up an inventory of these criticisms. In the empirical perspective under examination here, the problem is that this type of procedure is nothing more than a rhetorical device, insofar as the populations to which the test procedure is applied remain virtual in nature.

In practice, the researcher knows how to define their conditional variables on the basis of pairs: (smoker undergoing cessation, examination date) and (non-cessation smoker, examination date), assembled by her/him through observation. But what is the significance of the statistical population to which the inferential exercise makes reference? If we consider the *undergoing cessation* value, for example, how should the statistical population of the (smoker undergoing cessation/examination date) pairs be defined? Let us imagine a survey which would enable the anxiety score for all the human beings on the planet with the status of 'smoker undergoing smoking cessation' to be known on a certain date each month in the interval of time under consideration. We would then have as many populations as we have monthly surveys; we could then consider grouping together all of these monthly populations to define the population of observations relating to the 'cessation' status. There is not *one single* population, but rather *a number of* virtual populations. The null hypothesis is therefore based on a mental construct. As soon as this is defined more precisely, questions arise as to its plausibility and the interest of the test. Indeed, why should a survey supply an anxiety variable whose conditional means, subject to change, are identical?

Ultimately, it appears that the null hypothesis test constitutes a decision-making procedure with respect to the plausibility of a hypothesis devoid of any determined empirical meaning. The statistical inference used in the operationalization system is an odd way of settling the issue of generality: it involves deciding whether the difference between observed means may be generalised, even if the empirical meaning of this generality has not been established.

The myth of the average smoker

The difference between the two anxiety means may be interpreted as the difference between the degree of anxiety of the average smoker undergoing cessation and the degree of anxiety of the average non-cessation smoker, which poses two problems. Firstly, the discrete nature of the anxiety score contains a logical dead-end, i.e. the use of an impossibility to describe something which is possible. Let us assume an anxiety questionnaire comprising five items with answers scored 0, 1, 2 or 3, such that the score attributed to any group of 5 responses will fall within the sequence of natural numbers (0, 1, ..., 15). A mean score of 8.2 may indeed 'summarise' a set of scores, but cannot *exist* as an individual score. Consequently, should we wish to use a mean score to describe a typical smoker, it must be recognised that such a smoker is not possible and therefore not plausible. As a result, the difference between the two means cannot be used to describe the difference in degrees of anxiety of the typical smokers, unless it is admitted that a typical smoker is in fact a myth.

Let us now assume that the numerical coding technique enables a continuous variable to be defined by the use of so-called analogue response scales. The score of any smoker is *by definition* composed of the sum of two quantities, the mean score plus the deviation from the mean, the latter expressing the fact that the typical smoker is replaced in practice by a particular specimen of the statistical population, whose variable nature is *assumed* to be random – without it appearing necessary to have empirical grounds for the probability space on which this notion is based. In these conditions, the mean score constitutes a *parameter*, whose specification is an empirical matter inasmuch as the statistical population is actually defined. An empirical parameter is not, however, the same thing as an empirical law.

Formalisation of an empirical law

According to the nomothetic perspective, scientific ambition consists in discovering laws, i.e. general implications.[1] A general implication is a statement in the following form

$$\forall\, x \in A,\ p(x) \Rightarrow q(x), \tag{1}$$

which reads thus "for any x of A, if $p(x)$ then $q(x)$", where x is any component of a given set A, and $p(\cdot)$ and $q(\cdot)$ are singular statements. This formalisation applies without any difficulty to any situation in which the researcher has a pair of variables $(X,\ Y)$, from a domain $\Omega_n = \{\omega_i,\ i = 1,\ \ldots,\ n\}$, whose elements ω_i are pairs (person, observation date). The codomain of the independent variable X is a descriptive reference system of initial conditions $M(X) = \{x_i,\ i = 1,\ \ldots,\ k\}$, whilst the dependent variable, Y, specifies a value reference system, $M(Y) = \{y_i,\ i = 1,\ \ldots,\ l\}$, the effective observation of which depends, by hypothesis, on the independent conditions. Thus, the ontological substrate of an empirical law is the observation reference system $\Omega \times M(X) \times M(Y)$, where $\Omega \supset \Omega_n$ is an extrapolation of Ω_n: any element of Ω is, as a matter of principle, assigned a unique value in $M(X) \times M(Y)$ by means of the *function* $(X,\ Y)$.

Two comments arise from this definition. Firstly, as noted by Popper (1959), "[natural laws] do not assert that something exists or is the case; they deny it" (p. 48). In other words, they state a *general* ontological impossibility in terms of $\Omega \times M(X) \times M(Y)$: a law may indeed by formulated by identifying the initial conditions $\alpha(X) \subset M(X)$ for which a non-empty subset $\beta(Y) \subset M(Y)$ exists such that,

$$\forall\, \omega \in \Omega,\ X(\omega) \in \alpha(X) \Rightarrow Y(\omega) \in \beta(Y). \tag{2}$$

This formulation excludes the possibility of $X(\omega) \in \alpha(X)$ and $Y(\omega) \in \complement\beta(Y)$ being observed, where $\complement\beta(Y)$ designates the complementary set $\beta(Y)$ with respect to $M(Y)$. Making a statement in the form of (2) amounts to stating a general empirical fact in terms of Ω_n, and an empirical law in terms of Ω, by inductive generalisation. This law can be falsified, simply by exhibiting an example of what is said to be impossible in order to falsify it. The general nature of the statement stems from the quantifier $\forall$ and its empirical limit is found in the extension of Ω. The law may then be corroborated or falsified. If it is corroborated, it is possible to measure its degree of corroboration by the number of observations applying to it, i.e. by the cardinality of the equivalence class formed by the antecedents of $\alpha(X)$ – the class is noted $\mathrm{Cl}_{\Omega_n/X}[\alpha(X)]$.

[1]This is a more general and radical restatement of the definition given by Piaget (1970) of the notion of laws. For him laws designate "relatively constant quantitative relations which may be expressed in the form of mathematical functions", "general fact" or "ordinal relationships [...,] structural analyses, etc. which are expressed in ordinary language or in more or less formalised language (logic, etc.)" (p. 17).

The second comment relates to the notion of partial determinism. The mathematical culture passed on through secondary school teaching familiarises honest researchers with the notion of numerical functions $y = f(x)$, which express a deterministic law, i.e. that x being given, y necessarily has a *point* value. If the informative nature of the law is envisaged in negative terms (Dubois & Prade, 2003), the necessity of the point is defined as the impossibility of its complement. In the field of humanities (Granger, 1995), seeking total determinisms appears futile, but this does not imply that there is no general impossibility in $\Omega \times M(X) \times M(Y)$ and therefore no partial determinism. The fact that partial determinism may not have a utility value from the point of view of social or medical decision-making engineering has nothing to do with its fundamental scientific value. The subject of nomothetic research therefore appears in the form of a 'gap' in a descriptive reference system, this gap being theoretically interpreted as the effect of a general ontological impossibility. This is why in teaching, a methodology to support the nomothetic goal of training student researchers to 'search for the impossible' is called for.

How to seek the impossible

Discovery of a gap in the descriptive reference system involves the discovery of a general empirical fact, from which an empirical law is inferred by extending the set of observations Ω_n to an unknown phenomenological field $\Omega \supset \Omega_n$ (e.g. future events). A general empirical fact makes sense only with reference to the descriptive reference system $M(X) \times M(Y)$. Practically speaking, dependent and independent variables are multivariate. Let $\mathbf{X} = (X_1,\ X_2,\ \ldots,\ X_p)$ be a series of p independent variables and $M(\mathbf{X})$ the reference system of $\mathbf{X}$; $M(\mathbf{X})$ is the Cartesian product of the p reference systems $M(X_i)$, $i = 1,\ \cdots,\ p$. Similarly, let $\mathbf{Y} = (Y_1,\ \ldots,\ Y_q)$ be a series of q dependent variables and $M(\mathbf{Y})$ the reference system of $\mathbf{Y}$. The descriptive reference system of the study is therefore

$$M[(\mathbf{X},\ \mathbf{Y})] = M(\mathbf{X}) \times M(\mathbf{Y}) = M(X_1) \times M(X_2) \times \cdots \times M(X_p) \times M(Y_1) \times \cdots \times M(Y_q). \quad (3)$$

Thus the contingency table (the rows of which represent the multivariate values of $\mathbf{X}$, and the columns the multivariate values of $\mathbf{Y}$) can be defined. Observation readings are then carried out so that the cells in the contingency table are gradually filled in... or remain empty.

Two cases must be distinguished here. The first corresponds to the situation in which the researcher is completely ignorant of what is happening in their observation reference system, in other words, they does not have any prior observations. They therefore has to carry out some kind of survey in order to learn more. Knowing what is happening in the reference system means knowing the frequency of each possible state. It does not involve calling on the notion of probability (the latter being firmly in the realm of mathematical mythology) since it would involve knowing the *limit* of the frequency of each cell in the contingency table as the number of observations (n) tends towards *infinity*.

A *nomothetic gap* arises when there is at least one empty cell in at least one row of the contingency table, when the margin of the row (or rows) is well above the cardinality of $M(\mathbf{Y})$. It is possible to identify all the gaps in the reference system only if its cardinality is well below the cardinality of Ω_n, n. This empirical consideration sheds light on a specific epistemological drawback in Psychology: not only are its descriptive reference systems not given naturally, as emphasised by Danziger (1990),[2] but in addition the depth of constructible reality is such that its cardinality may be gigantic – so much so that discussing what is happening in an observation reference system cannot be achieved in terms of sensible intuition. The fact is that the socio-technical norms which shape the presentation of the observation techniques used in empirical studies do not refer either to the notion of descriptive reference system or the necessity of plotting the cardinality card$[M(\mathbf{X}) \times M(\mathbf{Y})]$ against the cardinality of the set of observations, card$(\Omega_n) = n$. If the quotient card$[M(\mathbf{X}) \times M(\mathbf{Y})]/n$ is not much lower than 1, planning to carry out an exhaustive examination of the nomothetic gaps in the descriptive reference system is unfeasible. This does not prevent the researcher from working on certain initial conditions $\alpha(\mathbf{X})$, but in such cases it must nonetheless be established that dividing the number of values of $M(\mathbf{Y})$ by the cardinality of the class $\mathrm{Cl}_{\Omega_n/\mathbf{X}}[\alpha(\mathbf{X})]$ of antecedents of $\alpha(\mathbf{X})$ in Ω_n gives a result which is far less than 1.

Let us now present the second case, for which it is assumed that the researcher has been lucky enough to observe the phenomenon of a gap, whose 'coordinates' in the descriptive reference system of the study are $[\alpha(\mathbf{X}),\ \complement\beta(\mathbf{Y})]$. The permanent nature of this gap constitutes a proper *general hypothesis*. This hypothesis should be *tested* using a *targeted* observation strategy. Indeed, accumulating observations in Ω is of interest from the point of view of the hypothesis if these observations are such that:

- $\mathbf{X}(\omega) \in \alpha(\mathbf{X})$, in which case we seek to verify that $\mathbf{Y}(\omega) \in \beta(\mathbf{Y})$,
- $\mathbf{Y}(\omega) \in \complement\beta(Y)$, in which case we seek to verify that $\mathbf{X}(\omega) \in \complement\alpha(\mathbf{X})$.

This approach to observation is targeted, and indeed makes sense, in that it focuses on a limited number of states: the researcher knows exactly what they is looking for. It is the very opposite of blindly reproducing an experimental plan or survey plan.

When a counterexample is discovered, i.e. ω_e exists such that $\mathbf{X}(\omega_e) \in \alpha(\mathbf{X})$ and $\mathbf{Y}(\omega_e) \in \complement\beta(\mathbf{Y})$, this observation falsifies the general hypothesis. The researcher can then decide either to reject the hypothesis or to defend it. If they decides to defend it, they may restrict the set of conditions $\alpha(\mathbf{X})$, or try to *find* a variable X_{p+1} which *modulates* verification of the rule. Formally speaking, this modulating variable is such that there is a strict non-empty subset of $M(X_{p+1})$ – let this be $\gamma(X_{p+1})$ – such that:

$$\forall\ \omega \in \Omega,\ [\mathbf{X}(\omega) \in \alpha(\mathbf{X}) \text{ and } X_{p+1}(\omega) \in \gamma(X_{p+1})] \Rightarrow \mathbf{Y}(\omega) \in \beta(\mathbf{Y}). \tag{4}$$

Irrespective of how they revises the original hypothesis, they will have to restrict its domain of validity with respect to the – implicit – set of possible descriptive reference systems. A major consequence of revising the law by expanding the descriptive reference system of initial conditions is resetting the corroboration counter, since the world being explored has been given an additional descriptive dimension: this is the reference system $\Omega \times M(\mathbf{X}_1) \times M(\mathbf{Y})$, where $\mathbf{X}_1 = (\mathbf{X},\ X_{p+1})$.

Example

Without it being necessary to develop the procedure presented here in its entirety, we can illustrate it using the example of smokers' anxiety. The problem consists of restating the

[2] "But in terms of truth, scientific psychology does not deal with natural objects. It deals with test scores, evaluation scales, response distributions, series lists, and countless other items which the researcher does not discover but rather constructs with great care. Conjectures about the world, whatever they may be, cannot escape from this universe of artefacts" (p. 2).

'general hypothesis' as a statement which is (i) general, properly speaking, as understood in (1) –, and (ii) falsifiable. We may proceed in two stages. Firstly, it is not necessary to talk in terms of reference systems to produce a general statement. Expressing the problem in terms of the difference between two means is not relevant to what is being sought; however, the idea according to which any smoker undergoing cessation *becomes* more anxious may be examined, along the lines of the 'general hypothesis' described by Fernandez and Catteeuw (2001). This idea is pre-referential inasmuch as we are unable to define a smoker, a smoker undergoing cessation, or a person who is becoming more anxious.

Since we cannot claim to be able to actually settle these issues of definition, we shall use certain definitions for the purposes of convenience. Let U be a population of people and T a population of dates on which they were observed. Let Ω_n be a subset of $U \times T \times T$ such that, for any triplet $\omega = (u,\ t_1,\ t_2)$, u is known on dates t_1 and t_2 in terms of their status as:

- a non-smoker, a smoker undergoing cessation or a non-cessation smoker
- and their state of anxiety, for instance with reference to a set of clinical signs, of which the person is asked to evaluate the intensity on date t, using a standard 'state-anxiety' questionnaire.

It can be noted that the set Ω_n is a finite, non-virtual set, in that a person u whose smoker status is not known on date t_1 or t_2 for example, constitutes a triplet which *does not* belong to this set. According to our approach to the statistical population, it is not necessary for the observations to be the result of applying a specific random sampling technique. Since Ω_n constitutes a set of known observations from the point of view of the descriptive reference system, it is a *numbered* set, to which new observations can be added over time; whence the notation $\Omega_{n_\jmath}$ (read "j-mat"), where $n_\jmath$ stands for the cardinality of the most recent update to the set of observations.

We can then define the following variables $\mathbf{X}_\jmath$ and $\mathbf{Y}_\jmath$, from the subset $P_\jmath$ of $\Omega_{n_\jmath}$, which includes the triplets $(u,\ t_1,\ t_2)$ such that $t_2 - t_1 = d$, where d is a transition time (e.g. 2 days). The variable $\mathbf{X}_\jmath$ matches any component of $P_\jmath$ with an image in $M(\mathbf{X}_\jmath) = \{nf,\ f_1,\ f_2\} \times \{nf,\ f_1,\ f_2\}$, where nf, f_1 and f_2 signify 'non-smoker', 'non-cessation smoker' and 'smoker undergoing cessation' respectively. Let us call $\alpha(\mathbf{X}_\jmath)$ the subset of $M(\mathbf{X}_\jmath)$ including all the pairs of values ending in f_2 which do not begin with f_2 and take an element $p \in P_\jmath$: the proposition '$\mathbf{X}_\jmath(p) \in \alpha(\mathbf{X}_\jmath)$' means that in the period during which they was observed, person u had been undergoing smoking cessation for two days whereas she/he has not been before.[3]

The dependent variable $\mathbf{Y}_\jmath$ must now be defined. Let us assume that for any sign of anxiety, we have a description on an ordinal scale (i.e., a Likert scale). Anxiety can then be described as a multivariate state varying within a descriptive reference system A. Consider $A \times A$; in this set a subset $\beta(\mathbf{Y}_\jmath)$ can be defined which includes changes in states defined as a worsening of the state of anxiety. The variable $\mathbf{Y}_\jmath$ can then be defined, which, for each $p \in P_\jmath$, corresponds to a state in $M(\mathbf{Y}_\jmath)$. The proposition '$\mathbf{Y}_\jmath(p) \in \beta(\mathbf{Y}_\jmath)$' signifies that in the period during which they was observed, person u became more anxious. Lastly, the

[3] It may be noted that an observation p such that $\mathbf{X}_\jmath(p) = (nf,\ f_2)$ is not plausible; this relates to the question of the definition of the state of cessation and does not affect the structure of the logic.

general hypothesis can be formulated in terms which ensure that it may be falsified:

$$\forall\, p \in P_\jmath,\ \mathbf{X}_\jmath(p) \in \alpha(\mathbf{X}_\jmath) \Rightarrow \mathbf{Y}_\jmath(p) \in \beta(\mathbf{Y}_\jmath). \tag{5}$$

We have just illustrated an apparently hypothetical-deductive approach; but in fact it is an exploratory procedure if the community is not aware of any database enabling a nomothetic gap to be identified. Let us assume that the work of the researcher leads to the provision of a database Ω_{236} for the community and that sets $\alpha(\mathbf{X}_\jmath)$ and $\beta(\mathbf{Y}_\jmath)$ are defined after the fact, such that at least one general fact may be stated. The community with an interest in the general fact revealed by this data may seek new supporting or falsifying observations in order to help update the database.

If a researcher finds an individual v, with $q = (v,\ t_{v1},\ t_{v2})$ and $t_{v2} - t_{v1} = d$, such that $\mathbf{X}_\jmath(q) \in \alpha(\mathbf{X}_\jmath)$ and $\mathbf{Y}_\jmath(q) \in \complement\beta(\mathbf{Y}_\jmath)$, this means that there is a smoker who has been undergoing cessation for two days, whose anxiety has not worsened. Let us assume that the researcher investigates whether the person was already very anxious; they may suggest that rule (5) should be revised so as to exclude people whose initial clinical state corresponds to certain values in the reference system A. This procedure usually consists in restricting the scope of validity of the general hypotheses.

Discussion

Operationalization in Psychology consists in restating a pre-referential proposition in order to enable the researcher to test a statistical null hypothesis, the rejection of which enables the 'general hypothesis' to be credited with a certain degree of acceptability.[4] Using an example taken from Fernandez and Catteeuw (2001), we have shown that the aim of such a procedure is not the discovery of empirical laws, i.e. the discovery of nomothetic gaps in a reference system. We shall discuss two consequences of our radical approach to seeking empirical laws in an observation reference system $\Omega \times M(\mathbf{X}) \times M(\mathbf{Y})$. The first relates to the methodology for updating the state of knowledge in a field of research, the second to the probabilistic interpretation of accumulated observations.

The state of knowledge in a given field of research can be apprehended in practical terms by means of a list of m so-called scientific publications. Let us call this set composed of specialist literature L_m and let l_i be an element in this list. The knowledge historian can then ask the following question: does text l_i allow an observation reference system of the type $\Omega_n \times M(\mathbf{X}) \times M(\mathbf{Y})$ to be defined? Such a question can only be answered in the affirmative if it is possible to specify the following:

1. $n > 0$ pairs $(u,\ t)$,
2. $p > 0$ reference systems enabling the description of the initial conditions affecting the n pairs $(u,\ t)$,
3. $q > 0$ reference systems enabling the description of states affecting the n pairs $(u,\ t)$ according to the initial conditions in which they are found.

Specifying a descriptive reference system consists in identifying a finite set of mutually exclusive values. Not all the description methods used in Psychology allow such a set to be defined; for example, a close examination of the so-called Exner scoring system (1995)

[4] Meehl (1967) noted several decades ago that the greater the 'experimental precision', i.e. sample size, the easier it is to corroborate the alternative hypothesis.

for verbatims which may be collected for any Rorschach (1921) test card did not enable us to determine the Cartesian product of the possible values. And yet, to find a gap in a reference system, this reference system must be constituted, so as to form a stabilised and objective descriptive framework. Faced with such a situation, a knowledge historian would be justified in describing a scientific era in which research is based on such a form of descriptive methodology as being a pre-referential age.

With regard to the matter of the objectivity of a descriptive reference system, we shall confine ourselves to introducing the notion of *score-objectivity.* Let $P = \{p_i,\ i = 1,\ \ldots,\ z\}$ be a set of Psychologists and $\omega_j \in \Omega$. $(\mathbf{X},\ \mathbf{Y})_i(\omega_j)$ is the value of ω_j in $M(\mathbf{X}) \times M(\mathbf{Y})$ as determined by the Psychologist p_i. We may say that $M(\mathbf{X}) \times M(\mathbf{Y})$ is score-objective relative to P if $(\mathbf{X},\ \mathbf{Y})_i(\omega_j)$ depends only on j for all values of j. If a descriptive reference system is not score-objective, an event in $\Omega \times M(\mathbf{X}) \times M(\mathbf{Y})$ which occurs in a gap cannot categorically be interpreted as a falsifying observation, since it may depend on a particular feature of the way the reporting Psychologist views it. Unless and until the descriptive definition of an event is regulated in a score-objective manner, the nomothetic aspiration appears to be premature, since it requires the objective world to be singular in nature.[5] Only once a descriptive reference system has been identified may the knowledge historian test its score-objectivity experimentally.

The historian might well discover that a field of research is in fact associated with the use of divergent description reference systems. Their task would then be to connect these different fields of reality by attempting to define the problem of the correspondence between the impossibilities identified in the field R_a and the impossibilities identified in the field R_b – which assumes such identification is possible. Given a certain descriptive reference system of cardinality c, the historian may evaluate its *explorability* and perhaps note that certain description reference systems are inexplorable. Concerning explorable reference systems, they could perhaps try to retrieve data collected during the course of empirical studies, constitute an updated database, and seek nomothetic gaps in it.[6]

Let us now move on to the second point of this discussion. If the reference system is explorable and assumed to be score-objective, it may be that each of its possible states has been observed at least once. In this case, the descriptive reference system is sterile from the nomothetic point of view and this constitutes a singular observation fact: everything is possible therein. In other words, given an object in a certain initial state, nothing can be asserted regarding its $\mathbf{Y}$-state. This does not prevent the decision-making engineer from wagering on the object's $\mathbf{Y}$-state based on the distribution of $\mathbf{Y}$-states, conditioned by the initial conditions in which the object is found. These frequencies may be used to measure 'expectancies', but they do not form a basis on which to *deduce* the existence of a probability function for these states. Indeed, defining a random variable $\mathbf{Y}$ or $\mathbf{Y}|\mathbf{X}$ requires the definition of a probability space on the basis of the possible states $M(\mathbf{X}) \times M(\mathbf{Y})$. In order to be probabilistic, such a space requires a probability space established on the basis of Ω (e.g.,

[5]We cannot simply classify the sources of score-subjectivity as measurement errors in the quantitative domain (Stigler, 1986), since most descriptive reference systems in Psychology are qualitative; diverging viewpoints for the same event described in a certain descriptive reference system represent an error, not of measurement, but of definition.

[6]This type of database, established by merging several databases, has nothing to do with the aggregation methodology of 'meta-analyses' based on the use of statistical summaries (e.g., Rosenthal & DiMatteo, 2001).

Renyi, 1966). Since Ω is a virtual set, adding objective probabilities to it is wishful thinking: seeing ($\mathbf{X}$, $\mathbf{Y}$) as a pair of random variables constitutes an unfalsifiable interpretation. Since such an interpretation is nonetheless of interest for making decisions, the existence of a related law of probability being postulated, the probability of a given state may be estimated on the basis of its frequency. The higher the total number of observations, the more accurate this estimation will be, which is why a database established by bringing together the existing databases is of interest. With the advent of the internet, recourse to probabilistic mythology no longer requires the inferential machinery of null-hypotheses testers to be deployed; it rather requires the empirical stabilisation of the parameters of the mythical law.

We conclude this critical analysis with a reminder that scientific research in Psychology is also aimed at the discovery of empirical laws. This requires two types of objectives to be distinguished with care: practical objectives, which focus on decision amid uncertainty, and nomothetic objectives, which focus on the detection of empirical impossibilities. Has so-called scientific Psychology been able to discover any empirical laws, and if so, what are they? From our contemporary standpoint, this question is easy to answer in principle – if not in practice.

References

Charbonneau, C. (1988). Problématique et hypothèses d'une recherche. In M. Robert (Ed.), *Fondements et étapes de la recherche scientifique en psychologie* (3e ed., pp. 59–77). Edisem.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, *30*, 116–127.

Danziger, K. (1990). *Constructing the subject: Historical origins of psychological research.* New York: Cambridge University Press.

Dubois, D., & Prade, H. (2003). Informations bipolaires: une introduction. *Information Interaction Intelligence*, *3*, 89–106.

Exner, J. E. Jr. (1995). *Le Rorschach : un système intégré.* Paris: Éditions Frison-Roche (A. Andronikof, traduction).

Fernandez, L., & Catteeuw, M. (2001). *La recherche en psychologie clinique.* Paris: Nathan Université.

Granger, G.-G. (1995). *La science et les sciences* (2e ed.). Paris: Presses Universitaires de France.

Krueger, J. (2001). Null hypothesis significance testing. *American Psychologist*, *56*, 16–26.

Meehl, P. H. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, *34*, 103–115.

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, *5*, 241–301.

Piaget, J. (1970). *Épistémologie des sciences de l'homme.* Paris: Gallimard.

Popper, K. R. (1959). *The logic of scientific discovery.* Oxford England: Basic Books.

Renyi, A. (1966). *Calcul des probabilités.* Paris: Dunod (C. Bloch, traduction).

Reuchlin, M. (1992). *Introduction à la recherche en psychologie.* Paris: Nathan Université.

Rorschach, H. (1921). *Psychodiagnostik.* Bern: Bircher (Hans Huber Verlag, 1942).

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, *52*, 59–82.

Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900.* Cambridge, MA: The Belknap Press of Harvard University Press.

Vautier, S. (2011). How to state general qualitative facts in pychology? *Quality and Quantity*, 1–8. URL http://dx.doi.org/10.1007/s11135-011-9502-5.