Could it be the case that if I am right my opponents will be pleased? A rejoinder to Johnson-Laird, Byrne and Girotto

Guy Politzer

Could it be the case that if I am right my opponents will be pleased ? A rejoinder to Johnson-Laird, Byrne and Girotto

Guy Politzer

Abstract

I take up the four issues considered by Johnson-Laird, Byrne and Girotto in their reply to Politzer (2007). Based on the conceptual clarification which they adduce, it seems that the disagreement can be settled about the first one (truth functionality) and can be attenuated about the second one (the paradoxes of material implication). However, I maintain and refine my criticisms on the last two (negation and the probability of conditionals), backed up by considerations borrowed from the perspective of the conditional probability semantics for conditionals.

Guy Politzer
CNRS – Institut Jean Nicod
Ecole normale supérieure
29 rue d'Ulm, 75005 Paris, France
email: politzer@univ-paris8.fr

1 Introduction

In their reply to my review and critical appraisal of the psychological theories of reasoning with conditionals (Politzer, 2007), Johnson-Laird, Byrne and Girotto (henceforth JLBG) consider four main issues. These are truth functionality, the paradoxes of material implication, negation and the probability of conditionals. I take up these points in turn and finish with some comments made from the perspective of the conditional probability semantics for conditionals.

2. Truth-functionality

I believe that the debate is relatively minor and stems from the lack of availability of a specific expression to label the mental model theory (MM theory). The meaning of a conditional sentence refers to a set of possibilities –in fact ten different sets after modulation (Johnson-Laird and Byrne, 2002). Because each of these sets can be mapped one-to-one on a truth table they can be described in terms of truth value (even though the input is not a truth value but an intensional representation). But the important point which should close the debate is that these sets of possibilities are not determined by compositionality (as truth-functionality would require) but by semantic and pragmatic modulation. It may, however, be worth pointing out that in case the operation of modulation leaves the core meaning unaltered, the possibilities of the "conditional interpretation" (one of the ten sets) can still be *described* in terms of truth value as a material conditional.

3. The paradoxes

JLBG write (p 10 of manuscript):

> if the conclusion of a valid inference throws information away by adding a
>
> disjunctive alternative to the possibilities consistent with the premises, then
>
> it should be odd, and individuals should balk at it. The challenge to critics is
>
> to find a counterexample.

At first sight, based on this formulation, a counterexample, which they challenge

critics to show, suggests itself, and this is modus ponens. The premises of MP

(A $\rightarrow$ C and A) are equivalent to A&C. The conclusion C is equivalent to

(A&C) v ($\neg\Box$&C), so that a rewriting of MP is: A&C / (A&C) v ($\neg\Box$&C) where the

second disjunct does throw out semantic information by adding a disjunct to the

single possibility consistent with the premise. This illustrates that there are

always at least as many *logical* possibilities in the conclusion as there are in the

premises of any valid deduction (and there are exactly as many in case of an

identity). But here is the origin of the debate: the possibilities that are

considered by JLBG are *psychological* possibilities (Johnson-Laird, personnal

communication), that is, mental models. The rewriting of modus ponens should

now be [a c] for the premise and [c] for the conclusion (without development)

which shows that no disjunct is added to the possibility consistent with the

premise. In brief, there is ambiguity in JLBG's use of "possibility"; after

disambiguation, my objection and counterexample do not hold any more.

There is, however, more to discuss about the paradoxes of thematerial

conditional. If JLBG's explanation of the paradoxes is right, two inferences that

can be described by the same sets of possibilities should produce similar estimates of acceptability. Consider the following inferences (from Johnson-Laird and Byrne, 2002):

There isn't a circle; therefore, if there is a circle then there is triangle

and its logically equivalent disjunctive formulation,

There isn't a circle; therefore, there is not a circle or there is triangle.

Are both equally rejected? If one considers only the forward inferences (when the conclusion is to be discovered or evaluated, as is usually the case in experiments), intuition seems to confirm the MM claim: both are odd. Now compare the *backward* inferences in dialogues D1 and D2:

(D1) Peter: *There isn't a circle or there is a triangle*

Mary (who has got evidence): *You are right: There isn't a circle.*

Mary has offered a convincing reason that justifies the truth of Peter's assertion and this is based on the validity of the argument, which is fully accepted.

(D2) Peter: *If there is a circle then there is a triangle*

Mary (who has got evidence): *You are right: There isn't a circle.*

Now Mary has offered an intuitively opaque and unconvincing reason, and her argument seems as odd as its forward counterpart. Notice that one cannot object some specificity pertaining to the backward inference because this should apply equally to the conditional and the disjunctive inferences. The result of this comparison is twofold: one, contrary to the MM claim, there is something specific to *if* in the oddity of the paradoxes, and two, the analysis of the paradoxes in terms of sets of possibilities offers a necessary but not sufficient condition.

4. The negation of conditional sentences

Why do individuals tend to deny conditional sentences *if A, C* by *if A, not C* rather than by *A and not C* (an expression of the impossible case)? JLBG have a straightforward explanation: Negation, like other operators, has a tendency to apply to the main clause, so that neg(A →C) is interpreted as A →neg C. So, individuals are incorrect because the resulting sentence and the initial one cannot contradict each other. In addition, it is claimed that a better task to reveal the correct negation is to ask what is impossible. I discuss these claims in turn.

The negation lowering explanation is worth considering as it is parcimonious and linguistically documented. But it answers only part of the question while failing to account for a basic psychological phenomenon. It proposes an explanation for the choice of *if A, not C,* but not for the near absence of *A and not C.* This is all the more puzzling as [A ¬C] is the explicit, and presumably more accessible model for *if A, not-C.* But even if this could be explained, the hypothesis does not account for linguistic intuition. When one considers how to deny a sentence such as *if it rains, the match will be canceled,* the sentence *if it rains, the match will not be canceled* meets the expectations of many people, whereas *it will rain and the match will not be canceled* sounds anomalous and seems to miss something semantically essential. This is the basic phenomenon which an adequate theory should explain and which the negation lowering hypothesis fails to do. An explanation will be proposed below.

Besides this shortcoming, the syntactic explanation is made at an exorbitant cost for the plausibility of MM theory. It is claimed that individuals

who negate a conditional sentence by *if A, not C* are incorrect because the resulting sentence has a possibility in common with the initial sentence (in fact, two possibilities, those when A is false), which means that the two sentences cannot contradict each other. It is uncontroversial that negation is a demanding operation. But could it be so demanding that individuals usually fail to correctly negate one of the most basic types of sentences they produce and understand in their daily life? Of course, individuals commit fallacies occasionally; but the error attributed to them by JLBG is systematic and, presumably, universal as everybody uses such denials in daily life. Individuals who routinely make incorrect negation are bound to be inconsistent, so that communication would usually fail. But this is not the case. As I will show below, individuals use negated conditionals to convey a whole range of nuances in a refined way.

JLBG claim that to reveal negation of *if A, C,* it is more appropriate to ask reasoners what is impossible given *if A, C* than to ask them to negate the sentence. The individuals' capability to represent *A and not C* as a counterexample to a conditional sentence *if A, C* is not a matter of debate: presumably all theories agree on this. Because this is what the "impossible" task tests, the results do not support the MM theory any more than other theories.

5. The probability of conditionals

Space now allows an explanation of why I found Girotto & Johnson-Laird's (2004) interpretation of their experiments on the probability of conditional

sentences debatable. One, there is a problem of validity. Participants face mathematical problems that belong to the domain of probability. Except for one problem in which the probability space is particularly simple, a combinatorial analysis is required before the ratio of probabilities is computed. This is a task in which participants must use technical skills or concepts of probability to succeed, far away from the intuitive judgment of how likely is a consequent given an antecedent for conditionals of daily life, so that what is tested is a confounding of mathematical skills and of the comprehension of the question. The high frequency of unexpected answers by whichever theory and the low level of performance on the first two problems clearly support this view. Two, the rationale for the conjunctive strategy is missing. No reason is offered why, when individuals are able to envisage all the possibilities consistent with the conditional, there are individuals who fail to treat this sentence as true in these possibilities (that is, why they fail to adopt the "complete" strategy). The data show hardly any evidence of the operation of the "complete" strategy (even though it is motivated) but the conjunctive strategy (that is unmotivated) occurs one third of the time on the average.

The major current approach to the probability of the conditionals is based on the notion, fully developed by Adams (1975), that the probability of a conditional sentence is the conditional probability of its consequent on its antecedent ($pC \mid A$, the PCasCP hypothesis). JLBG's have three misgivings with the PCasCP hypothesis. One concerns the paraphrase of the conditional with a disjunction (their probabilities differ). They take for granted individuals' capacity to paraphrase *if A, C* with *not-A or C* on the basis of the results of Richardson

and Ormerod (1997). But the methodology is flawed for, in all likelihood, participants learnt the experimenters' transparent objective through repeated measures. In addition, the problem is far more subtle because probabilistic logics predict that it is valid to infer from *if A, C* to *not-A or C,* but not always in the reverse direction, and crucially, the validity of the latter depends on the specific values of pA and pC. So, carefully controlled experiments are required. There are, however, seminal observations made by Fillenbaum (1978) using a free production task. This showed participants' general low level of paraphrases by *or* (possibly due to the lack of examples in the instructions, a defect in the opposite direction). But there is an interesting result: there was a notable exception for two specific relations (logical truths and threats). As noted by JLBG, p($\neg$AvC) differs from pC$\vert$A: the latter is smaller, except in the case of certainty where both equal 1. Now, it is remarkable that these two cases where the numerical equivalence is uncontroversial are precisely those where the conditional aims to convey full belief, whether in the expression of a logical truth or in the expression of a threat; in the latter case the certainty for the hearer that the consequent will be executed by the speaker is essential for the speech act to have perlocutionary force.

Two, in support of their view they accept the results of Schroyens, Schaeken and Dieussaert (2008) showing that the material conditional interpretation can be boosted by appropriate priming manipulations: attracting the attention to the possibility of the non-A cases results in greater chance that these cases will be incorporated in the calculation of the probability of the statement's truth, leading to the material conditional. But this result does not

constitute evidence against a probabilistic view. Naïve reasoners do express conditionals that converge with the material conditional. The sentence *if these two numbers are consecutive, their sum is odd* expresses a mathematical truth uttered with certainty ($pC \mid A = 1$). It is a limiting case where the probabilistic import dissolves into certainty and which can receive a material conditional interpretation. The material conditional is teacheable to some extent (painstakingly, as math and logic teachers experience daily); it may become more or les available in the individual's metacognitive equipment, depending on the kind of formal education they receive. Training can also be attempted in an experimental setting but this is a hard enterprise, like in the Schroyens et al. study where the effect size amounted to an increase of about 10% and did not even reach significance in the second experiment.

Three, is conditional probability at the border of individuals' competence? They often err in tasks that require mathematical computation, that is, they have difficulty understanding the formal concept. In daily life, where conditional probabilities are pervasive, there are two common kinds of mistake. One is the confusion of $pY \mid X$ with pXY, which the formula explains: this may occur when, for pragmatic reasons, the "sample space" (X) is identified with the whole universe. The other is the confusion, mentioned by JLBG, between $pY \mid X$ and $pX \mid Y$ which is well documented (e.g., Sherman et al., 1992) and again accounted for by the formulas $pY \mid X = pXY / pX$, and $p \, X \mid Y = pXY / pY$, showing that crucially they share the common term pXY but differ by their sample spaces. Now context and/or saliency may suggest focusing on one sample or the other, a potential source of mistake akin to figure-ground effect.

This is not specific to probability but affects also the expressions of frequency and proportion. For instance, I have repeatedly observed in classes of statistics that a phrase such as "the proportion of girls in the Science stream" is deeply ambiguous. It tends to be interpreted relatively more often as the ratio of girls to all the students in the Science stream (implicit contrast: boys) or as the ratio of female Science students to all female students (implicit contrast: non-Science stream), depending on how the topic is manipulated, e. g., gender versus professional prospect. In brief, (i) here again natural language is imperfect to convey mathematical --as well as logical-- concepts, and (ii) the formal, metacognitive concept is hard. But it would be a deep epistemological mistake to reject (or accept) the theorist's modelling on the basis of naïve subject's difficulty (or ease) in understanding the concepts and calculations that constitute the theorist's abstract model: these are at two different levels of analysis. (This remark also applies to the conclusions one might wish to draw, in whichever direction, from the Schroyens et al. study).

6. The probabilistic view

Politzer's review (2007) focused essentially on the two approaches to deduction that have been the most influential and also the most debated since the late 1970s (mental rules and mental models). They were developed to investigate individuals' performance in formal reasoning, which many thought to be the main, if not unique, test of human rationality. In recent years, there has been a major shift of interest towards tasks ecologically valid that take into account the

uncertainty of the premises. This basic characteristic of human reasoning was entirely neglected and only recently has a new look on conditionals been adopted by some researchers (Liu, Lo and Wu, 1996, Oaksford, Chater, and Larkin, 2000; Evans, Handley and Over, 2003). The most recent psychological theory of conditionals (Evans and Over, 2004) adopts the Ramsey test as a psychological operation on which the meaning of *if* is based. The PCasCP hypothesis follows naturally from this view and so does the explanation of individuals' intuitions that have been mentioned earlier regarding the denial of conditionals.

One major advantage of the PCasCP view is that it allows the representation of the uncertain status of the sentence and conveys its value. Now assuming that a speaker has asserted *if A, C* with full belief (pC│A = 1), her interlocutor may wish to express disagreement based on various epistemic states. This does not aim to convey a *logical* negation (a contradiction): rather this challenges the *assertability* of the conditional sentence. Total disagreement (pC│A = 0) or strong disagreement (pC│A low) can be expressed by *if A, not C,* which vindicates the correctness of negation lowering. But there are other nuances in the speaker's degree of belief that support an expression of disagreement: She may wish to convey only that p not-C│A > pC│A; this also can be expressed by *if A, not C* (but can be made more explicit by use of a modal, such as *it is likely that if A, not C).* Frequently, one may wish to convey a slight doubt about the conditional, that is, to alter belief from pC│A = 1 down to a value slightly lower than 1. This is typically captured by modal expressions such as *it is possible that if A, not C,* or *it is not certain that if A, C;* as a variant,

one may wish to convey a reason for such a reduced belief, which is typically captured by unconditional modal expressions of the type *it is possible that A and not C.* In brief, the probabilistic view explains the use of common verbal expressions to deny conditionals, including the critical *if A, not-C.* The latter, far from being a systematic fallacy, is correctly understood by the interlocutors as a *conditional denial,* that is, an expression of dissent—itself uncertain--about the likelihood of C under the supposition that A holds, in other words, about the assertability of *if A, C.*

Recently, Handley, Evans and Thompson (2006) asked participants presented with negated basic conditionals to judge what follows, what is possible and what is necessary; the results unambiguously supported the conditional probability prediction that for individuals nothing necessarily follows from a negated conditional. But supporters of MM theory might argue that the negated conditionals could not be interpreted as such because negation was given wide scope, so that the two approaches make the same predictions. In an unpublished pilot study, I have asked participants to deny future indicative conditionals such as *if it rains, the match will be canceled,* or *if it's cloudy, it will rain.* A speaker was supposed to have uttered the conditional sentence and the participants were asked to complete the dialogue by writing up the interlocutor's plausible reply prompted by "No...". By using this cue, the procedure invited participants to mentally deny the statement, which allowed them to generate the denial in their own words, without being imposed any syntactic construction. Importantly, they were not presented with a negated conditional sentence, which avoids the negation lowering objection. There was a rich variety of

expressions, but most of them could be classified into three groups of roughly equal frequency: One conditional, *if A, not C;* one modalised conditional: *if A, it is possible that not-C/not necessary that C;* and one modalised nonconditional (*sometimes/ it is possible that A and not C*). Significantly, there was not a single occurrence of categorical *A and not C.* Clearly, one cannot deny a *conditional* sentence by asserting a *categorical* sentence *A and not C*. This is because the speaker would forsake the suppositional import of the conditional, and this lies at the heart of individuals' linguistic intuition against such denials. There is one exception to this, namely when the speaker has knowledge that A has actually occurred (while C did not) in which case the suppositional status of A vanishes. This is well accounted for by the conditional probability view since a categorical assertion of *A & not-C* means p(A & not-C) = 1. But together with p not-C│A = 1 this requires pA = 1. Notice that this does not occur in case of a modal assertion of *A & not-C* which preserves the supposition by keeping pA < 1. In brief, uncertainty is pervasive: individuals generally reason with assumptions that have various degrees of belief and this applies to conditionals in the first place.

7. Conclusion

In all likelihood, the question I raised in the title will be answered negatively by most individuals, that is, people (and presumably my opponents too) believe that *if I am right, my opponents will not be pleased*. If I were unwise enough to think that I have an apodeictic case, I could express a denial based on such a thought by *I am right and my opponents will be not be pleased*. Surprisingly, this

is also what, according to their theory, my opponents themselves must believe. But, obviously, this yields much more than what one expects their opponents to be willing to envisage or concede, which is only that *if I am right they will not be pleased*, without removing the suppositional import of the conditional.

As illustrated in sections 5 and 6, conditional sentences have an inherent probabilistic import. In addition, communication and inference is often, if not generally, carried out about uncertain assumptions (which includes the conditional's components). It follows that an adequate theory of conditionals – whether at the semantic or the inferential level— should incorporate uncertainty, that is, have a means to represent degrees of belief and to explain how these propagate.

The MM theory constitutes a great intellectual achievement. The part of the theory that deals with conditionals contains many deep insights, but from the viewpoint just mentioned it does not meet the expectations one might have. *In its current state* of development, it is basically focused on a bi-valued view of reasoning, whether in terms of possibilities or in terms of truth. The theory may, in principle, have the tools to develop a model of reasoning with uncertain premises (especially for deduction) that could apply also to conditionals. Johnson-Laird et al. (1999) state that models can be tagged with numerical probabilities but this proposal has not been exploited so far, nor has its feasibility or psychological plausibility established: let us hope to see these developments in the future.

References

Adams E (1975) The logic of conditionals.Reidel, Dordrecht

Evans J St BT, Handley SJ, Over DE (2003) Conditionals and conditional probability. J Exp Psychol Learn Mem Cognit 29:321-335

Evans J St BT, Over, DE (2004) If. Oxford University Press, Oxford

Fillenbaum S (1978) How to do some things with *if.* In: Cotton J W & Klatzky RL (eds) Semantic factors in cognition. Lawrence Erlbaum Hillsdale, NJ, pp 169-214

Girotto V, Johnson-Laird PN (2004) The probability of conditionals. Psychologia 47:207-225

Handley SJ, Evans J St BT, Thompson VA (2006) The negated conditional: A litmus test for the suppositonal conditional? J Exp Psychol Learn Mem Cognit 32:559-569

Johnson-Laird PN, Byrne RMJ (2002) Conditionals: a theory of meaning, pragmatics, and inference. Psychol Rev 109:646-678

Johnson-Laird PN , Legrenzi P , Girotto V , Sonino-Legrenzi M , Caverni J-P (1999) Naive probability: A mental model theory of extensional reasoning. Psychol Rev 106:62-88

Liu I-M. , Lo K-C , Wu J-T (1996) A probabilistic interpretation of "If-Then". Quart J Exp Psychol 49A:828-844

Oaksford M, Chater N, Larkin J (2000) Probabilities and polarity biases in conditional inference. J Exp Psychol Learn Mem Cognit 26: 883-899

Politzer G (2007) Reasoning with conditionals. Topoi 26:76-95

Richardson, J Ormerod TC (1997) Rephrasing between disjunctives and
conditionals: Mental models and the effects of thematic content. Quart J Exp
Psychol 50A:358-385

Schroyens W, Schaeken W, Dieussaert C (2008) "The" interpretation(s) of
conditionals. Exp Psychol *55*:173-181

Sherman SJ, McMullen MN, Gavanski I (1992) Natural sample spaces and the
inversion of conditional judgments. J Exp Soc Psychol 28:401-421