190204

Guilt by Association: On Iffy Propositions and the Proper Treatment of Mental-Models Theory

Walter Schroyens and Walter Schaeken

Electronic version
URL: http://journals.openedition.org/cpl/411
ISSN: 1379-6100

Publisher
Centre PsyCLÉ

Electronic reference

This text was automatically generated on 19 April 2019.

© All rights reserved
Guilt by Association: On Iffy Propositions and the Proper Treatment of Mental-Models Theory

Walter Schroyens and Walter Schaeken

Introduction

The research on thinking and reasoning about conditionals has known a long history fuelled by the interest of cognitive scientists focussing on the linguistic, logical and psychological aspects of ‘if’. Such interest is readily framed by the centrality of conditional relationships. Our ability to reflect on potential implications of what is, could have been, or will be predicates some of the prime examples of human cognition. Evans et al. (2003) recently presented the results of three studies within a novel experimental paradigm. In this sentence-evaluation task people are informed about the exact number of cases of each of the four possible combinations between the affirmation or denial of antecedent (e.g. “The figure is red” or “The figure is blue”) and consequent (e.g., “The figure is a triangle, or ‘the figure is a Square’) of a conditional (e.g., ‘if the figure is Red, then it is a triangle’). People are then asked to judge how likely it is that the conditional is true of a particular case drawn at random form the set (with, i.c., a particular number of red or blue triangles and red or blue squares). Evans et al. (2003) distinguished three hypotheses about the interpretation of conditionals (see Table 1). These conjunctive, conditional-probability, and material-implication interpretations are defined as a function of the contingencies/possibilities people would represent/consider. Evans et al. (2003) showed that their participants’ behaviour is best explained by the conjunctive and the conditional-probability hypothesis, whereas their findings counter the material-implication hypothesis (also see Oberauer & Wilhelm, 2003). For instance, they showed that an increase in the absolute frequency of false-antecedent cases (e.g. 'Blue squares')}
vis-a-vis ‘If the figure is Red, then it is a Triangle’) reduced the truthfulness ratings of the conditionals, whereas one would expect the opposite under the material-implication hypothesis because such false-antecedent cases make the Material implication true (see Table 1). Evans et al.’s also regressed the observed truthfulness ratings to the predicted truth-fullness ratings and observed that individuals fell into two groups, whose performance is best explained by an interpretation that corresponds to the conjunction or the conditional probability. In Evans et al.’s discussion of the data they consider the mental-models theory. We show that the presumed evidence against material implication is not inconsistent with mental-models theory and that neither the conjunction hypothesis nor the conditional-probability hypothesis conflicts with a proper treatment of mental-models theory (i.e., Johnson-Laird & Byrne, 2002).

Table 1: Contingencies represented under different interpretations of a conditional of the form ‘if p then q’.

<table>
<thead>
<tr>
<th>Contingency</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>p→q</td>
</tr>
<tr>
<td>TF</td>
<td>p→¬q</td>
</tr>
<tr>
<td>FT</td>
<td>p→q</td>
</tr>
<tr>
<td>FF</td>
<td>p→¬q</td>
</tr>
</tbody>
</table>

Note. ‘TT’, ‘TF’, ‘FT’, ‘FF’ stands for contingencies wherein the antecedent/consequent is True or False.

Guilt by Association

Evans et al claimed to have “shown why the assumption of material implication embedded in the mental model theory of conditional reasoning (Johnson-Laird and Byrne, 1991, 2002) is unjustified” (p. 334). They draw this conclusion on the basis of their evidence showing that people do not represent ordinary conditional statements as material conditionals. We do not contest that in the context of Evans et al.’s experiments, ordinary conditional statements are not represented as material conditionals. In the following we will however expose the sophistry in the presumed implications of this finding for mental-models theory, which rests on misconceptions about the material-implication hypothesis and/or mental-models theory.

Evans et al associate the material-implication hypothesis with mental-models theory, which they can justifiably do within certain boundaries (see, the core meaning principle: Johnson-Laird & Byrne, 2002, p. 650). They claim that “under the material conditional, the probability of a conditional [if p then q] and its contra positive [if not-q then not-p] must be the same.” (Evans et al., 2003, p. 325). Given that it is easy to show that this ‘equivalence assumption’ is false, which they did by showing that the probability ratings of the conditional and its contrapositive differ substantially, they succumb to making mental-models theory guilty by association. Their guilt-by-association argument goes as
follows. First, they associate the material-implication hypothesis with mental-models theory (a psychological theory). Second, they then forget about psychology and make the logical equivalence-prediction. Third, they then show that the (logical) equivalence-hypothesis is false and argue that therefore the material-implication hypothesis (and a-fortiori) mental-models theory is (psychologically) false/invalid. The above argument attests to a form of so-called naive falsificationism1 and makes an unfair move in the evaluation of mental-models theory. The equivalence prediction is a logical and hence not necessarily a psychologically plausible prediction. It is misleading and false to suggest that the mental-models theory makes the equivalence assumption.

First, since the mental-models theory adheres to the implicit-model principle (Johnson-Laird & Byrne, 2002) it is false to claim or even suggest that the theory subscribes to the equivalence hypothesis. The implicit-model principle yields the following initial representation for ‘if p then q’:

\[
\begin{array}{cc}
p & q \\
\end{array}
\]

... 

The initial model for the contrapositive ‘if not-q then not-p’ is:

\[
\begin{array}{cc}
\text{not-q} & \text{not-p} \\
\end{array}
\]

... 

The non-equivalence in the initial-model representations clearly shows that the equivalence prediction cannot be made within mental-models theory. Second, it is misleading to suggest otherwise. How would a theory ever become “one of the most influential theories of human reasoning” (Evans et al., 2003, p. 324) if it cannot even account for the basic phenomena? For instance, Pollard and Evans (1980) asked whether the contrapositive follows from the conditional and observed that only 59% of the participants accepted the equivalence (if p then q, hence if not-q then not-p). They obviously did not give the option to indicate that ‘if p then q’ follows from ‘if p then q’.

Indeed, it has also long been established that a knowledge-lean Modus Ponendo Ponens argument (MPP: If p then q, p; therefore q) is almost universally accepted (Schroyens, Schaeken, & d’Ydewalle, 2001a, 2001b; also see Schroyens & Schaeken, 2003). It is also known that the Modus Tollendo Tollens argument (MTT: If p then q, not-q; therefore not-p) is endorsed much less frequently than MPP. If the equivalence hypothesis were to hold, then mental-models theory would not be able to account for the lower endorsement rates of MTT as compared to the near universally accepted MPP: the categorical premise ‘not-q’ in MTT defines an MPP on the contrapositive ‘if not-q then not-p’.

Not only mental-models theory rejects the equivalence assumption. None of the so-called mental-logics theories (see, e.g., Braine & O’Brien, 1998; Rips, 1994) subscribe to the equivalence hypothesis; and neither does the probabilistic model of Oaksford, Chater and Larkin (2000) for the obvious reason that the conditional probability of q, given p (the probability of MPP in this model) does not equal the conditional probability of not-p given not-q (i.e., the probability of MTT). Evans et al. are right in stating that “if p then q” is logically equivalent with “if not-q then not-p”. But this logical fact (a strawman) is not particularly pertinent since they are reporting/considering a psychological study/theory. In making the equivalence prediction one fails to distinguish logic and psychology. Evans et al. wrench several other logical equivalences. For instance, they state that “under the material conditional, the higher this [false-antecedent] probability, the higher should be the overall judged probability because material conditionals are always true when the
antecedent is false” (p. 13). That a material implication is true when the antecedent is false is a matter of textual fact in propositional calculus, but it is not what is considered to be the case in human reasoning. Evans et al (2003) also go on a slippery slope when stating that “thus according to Johnson-Laird and Byrne, the core meaning of an ordinary conditional makes it equivalent to the material conditional, \( \text{to either not-}p \text{ or } q \)” (p. 324). The equivalence between ‘if \( p \) then \( q \)’ and ‘either not-\( p \) or \( q \)’ is a formal logical equivalence; it is not supposed to be a psychological equivalence (Schaeken, Johnson-Laird, Byrne & d’Ydewalle, 1996). By not making a careful enough distinction between the principles of standard logic and the principles of the psychological mental-models theory, they have constructed unseemly arguments that make mental-models theory guilty by association.

“The” interpretation(s) of “if” in mental-models theory

A conjunctive interpretation of “if”

Under the conjunction-hypothesis people would consider only the contingency explicitly mentioned in ‘if \( p \) then \( q \)’. Evans et al. indeed observed increased ratings of the conditional’s probability when the frequency of the TT,[p,q] contingency was higher. They also observed a ‘false antecedent effect’: decreased probability ratings with an increase of false-antecedent cases. As a consequence of the set up of their experiments, the relative frequency of [p,q] cases decreased when the frequency of false-antecedent cases increased. The ‘false-antecedent effect’, thus, also concurs with a conjunctive interpretation. Evans et al. claim that it is “difficult to reconcile this finding with the claim of Johnson-Laird and Byrne (2002) that the core meaning of a conditional is to allow for the extension of the possibilities of \( p \text{, } q \), \( \sim p \sim q \text{, } \sim p q \) (or the equivalent TT, FT and FF cases)” (p. 329). When we gauge some of claims they made at other points, it becomes clear that the false-antecedent effect is not difficult to reconcile with mental models theory and was actually to be expected. Evans et al. themselves note that the false-antecedent effect is consistent with the conjunction hypothesis and at the same time mention that “Johnson-Laird and Byrne do predict that many people will give \( P(p.q) \) [i.e., the conjunctive probability] as the initial probability of the conditional” (p. 325; also see, Oberauer et al., 2003). It follows necessarily that Evans et al. have to agree that the false-antecedent effect is consistent with (if not predicted by) mental-models theory.

A probabilistic interpretation/evaluation of “if”

Evans et al. (2003) state that “[mental-models theory] is inconsistent with a relation between ordinary conditionals and conditional probability” (p. 324). There are multiple ways in which mental-models theory is consistent with conditional probability. Indeed, as noted already by Johnson-Laird and Byrne, the probabilistic interpretation ‘if \( p \) then probably \( q \)’ is not to be confused with the probabilistic evaluation ‘probably if \( p \) then \( q \)’. In both cases people would have a probabilistic representation of the two different possibilities (i.e. TT, [p,q] and TF, [p,not-q]), and a comparison of the frequency/probabilities yields a conditional-probability estimate. The presumed inconsistency of mental-models theory (Johnson-Laird & Byrne, 2002) rests on a uncharitable reading of two explanatory principles: the core-meaning principle and the truth principle. Indeed, Evans et al. also claimed to have “shown ... why the principle of truth in this theory is untenable” (p. 334). Consider the conditional:

“If the figure is a square, then the figure is red.”

about a set with 60 red and 10 yellow squares. The conditional cannot be true of the set and numerous studies (e.g., Evans et al., 1996) have shown that most people judge TF,
cases to be inconsistent or false. Given that people have the ‘Gricean’ tendency to assume (at least to start with) that the information one is confronted with is true (see, e.g., Grice, 1975; Johnson-Laird & Byrne, 2002; Johnson-Laird & Savary, 1999; Levinson, 2000; Schroyens, 1997, Schroyens, Schaken, & d'Ydewalle, 1996, 1999), they know, on the one hand, that it would not be possible that there are yellow squares. On the other hand they also see that there are in fact yellow squares. People have three options to cope with this inconsistency. First, they could question and/or deny the given facts. Second, they could question and modify their reading of the strict conditional ‘if p then q’ to the tautological, probabilistic conditional ‘if then possibly q’ or, third, they could accept the falsity of the conditional and evaluate ‘probably if p then q’. The latter two options seem to require the least cognitive effort (see, Sperber & Wilson, 1986).

When being faced with a tautological conditional or having such an interpretation of the conditional, people would form a representation of the following two possibilities:

square (red)
square (not-red)

... Indeed, “conditionals have mental models representing the possibilities in which their antecedents are satisfied” (Johnson-Laird and Byrne, 2002, p. 654). People would have little to no difficulty to elaborate their representation of the conditional with the explicitly provided frequentist information about the different possibilities represented in this initial mental-model set (see, Johnson-Laird, 1983; Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999): “If the cardinality of the set matters, then models can be tagged with numerals, just as they can be tagged to represent numerical probabilities (see, Johnson-Laird et al., 1999)” (Johnson-Laird & Byrne, 2003, p. 655). The numerals obviously matter when asked to rate the probabilities. That is, people would represent, in some way:

square-(red): 60
square-(yellow): 10

... This means that people would have little to no difficulty in recognizing that there are 70 instances where the figure is a square, and that there are 60 out of these 70 cases wherein this square is coloured red. In conclusion, mental-models theory is not inconsistent with a relation between the ordinary conditional and conditional probability, contrary to Evans et al. intuition. Conditional-probability estimates can arise from a tautological/probabilistic interpretation of ‘if’. Such a probabilistic interpretation is evidenced by a number of studies (see, e.g., Liu, Lo, and Wu, 1996; Oaksford et al., 2000; Stevenson & Over, 1995).

Instead of eliminating the inconsistency between the Gricean assumption that the utterances are true and the factual observation to the contrary by pragmatically modulating the proposition, people can simply abandon the Gricean assumption. People accept that the conditional is strictly false. Taking this option is more plausible in the context of Evans et al.’s (2003) experiments since it confirms the presumption inherent in questioning how likely the rule is true of a card from the set: when one considers how likely it is true, one presumes it is not true as such. In evaluating ‘probably/sometimes if p then q’, people would represent both the truth and falsity of the rule: Sometimes it is
true, sometimes it is false. In short, (some but not all) people would construct an initial representation of the conditional that also includes the truth-status of these TF cases.

This representation concurs with Ramsey's (1929/1990, p. 155) suggestion that:

If two people are arguing 'If p will q?' and both are in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; so that in a sense 'If p, q' and 'If p, not q' are contradictories.

The applicability of the Ramsey test is not surprising, since the test is explicitly intended for an evaluation of the question “if p will q?” and this is what Evans et al. (2003) asked their participants to do.

A probabilistic evaluation is not inconsistent with mental-models theory, and especially its truth-principle: “The principle does not imply, however, that individuals never represent false cases” (Johnson-Laird & Byrne, 2002, p. 654). When asked to contemplate on the falsity of something, people will do so. This does not impinge on the idea that people generally work under the assumption that the information they work with is true and will not consider the falsity of it without due reason. In presenting the truth-principle Johnson-Laird and Byrne (2002) state that “a mental model of an assertion represents a possibility given the truth of the assertion” (p. 653, italics added). The idea that people represent only true possibilities is indeed erroneous when one forgets that these possibilities would be presented given the truth of the assertion. It is actually this Gricean truth-assumption that is put into question when people are asked whether the conditional is true/false. It is no longer a given. The truth-principle hinges on the truth-assumption, the assumption that people generally adhere to the Gricean maxim of conversation that one’s contribution to discourse should be truthful (which also makes sense in terms of cognitive economy given the ontological primacy of ‘being’ vs. ‘not being’, truth vs. falsity/non-truth).

Conclusion

We have shown that mental-models theory is much more complex and with many more intricate details than would be apparent from Evans et al. treatment of it. Evans et al. focussed on the core-meaning assumption (Johnson-Laird & Byrne, 2002) and its undeniable link with the thesis that the sentential connective ‘if’ is the natural language equivalent of material implication in formal propositional calculus. This focus resulted in a neglect of other aspects of the theory and seems to have yielded the general idea that there is only one, a conditional/material-implication interpretation of the conditional. This made them claim that alternative interpretations are inconsistent with the theory.

Though Evans et al.’s data are interesting, their theoretical interpretation of the data as a critique of mental-models theory goes amiss.

In response to a more elaborate presentation of our critique (Schroyens & Schaeken, 2003), Evans (personal communication, June 2003) replies that:

"If the model theory gives up the principle of truth and then adds representation of frequencies plus a previously unspecified mechanisms of relative frequency judgement, then of course it can account for judgements of conditional probability theory. This, however, is not remotely the theory of JLB to which we refer in the paper"

This provides strong support for the validity of our arguments; the remaining point of contention being whether this is part of the theory as it is narrated by Johnson-Laird and Byrne (2002). It is misleading to say that it is not. Johnson-Laird and Byrne (2002)
obviously enough did not predict the different evaluation-types: Evans et al. (2003) did
construct a novel task. The fact that Johnson-Laird and Byrne (2002) did not make some
particular prediction and have not yet specified the predictions for new studies that
might become available in the future, however, does no imply that conjunctive and/or
conditional-probability evaluations are not predictable with the mental-models theory as
it was specified by Johnson-Laird and Byrne (2002). We will show that are indeed
predictable from Johnson-Laird and Byrne (2003).

In Evans' comment we quoted above, he repeats the three explanatory assumptions we
invoked. The textual evidence proofs that these are clearly not beyond Johnson-Laird and
Byrne (2002). First, "The principle does not imply, however, that individuals never
represent false cases" (Johnson-Laird & Byrne, 2002, 654). That is, one does not need to
abandon the truth-principle. Second: "if the cardinality of the set matters, then models
can be tagged with numerals, just as they can be tagged to represent numerical
probabilities (see Johnson-Laird et al., 1999)."(Johnson-Laird and Byrne, 2002, p. 655) That
is, it is not an ad hoc elaboration of ours to assume that frequencies can be represented.
Third, it seems rather obvious that new tasks require to specification of some new task-
specific mechanisms. Evans et al. (2003) were not impeded by the need for (and lack of) a
mechanism of relative frequency judgements when they discuss the material implication
hypothesis, which they associate with mental models theory: "... summing up the
probabilities of these models leads, of course, to a judgement equal to the probability of
the material conditional' (p. 325, italics added). Indeed, of course. In short, the expected
theoretical import of our argumentation – showing that it is false to claim that mental-
models theory "is inconsistent with a relation between ordinary conditionals and
conditional probability", by showing that it is not inconsistent – is agreed upon by J.
Evans.

Evans et al.'s unfortunate representation of mental-models theory appears to be due to
some misapprehension in their reading of Johnson-Laird and Byrne (2002). Evans et al.
repeatedly and consistently refer to the "core meaning of ordinary conditionals". Upon a
critical, close and careful reading of Johnson-Laird and Byrne (2002) this is almost a
contradiction in terms. Johnson-Laird and Byrne speak of the core meaning of basic
conditionals: "those [conditionals] with a neutral content that is as independent as
possible from context and background knowledge, and which have an antecedent and
consequent that are semantically independent apart from their occurrence in the same
conditional" (p. 649). Ordinary conditionals clearly are not, cannot be basic
conditionals. Ordinary conditionals are characterised by the fact that they are not neutral
in content, that they are dependent from context and background knowledge, and that
they have an antecedent and consequent that are not semantically independent. When
people are explicitly given information about the prevalence of particular contingencies,
as was the case in Evans et al. (2003), we are not dealing with a basic conditional. The
conditional “if the card is yellow, then it has a circle printed on it”, is clearly not
independent from a context that gives the following information about a deck of cards:

There is 1 yellow circle,
There are 4 yellow diamonds,
There are 16 red circles
There are 16 red diamonds

About this deck of cards (with a particular frequency distribution), people are asked to
judge how likely it is that the conditional is true of a card randomly drawn at random
from it. Asking this about the deck of cards ensures that the contextual information becomes relevant, and ascertains that we are not dealing with a basic conditional that is neutral in content and independent of the context. Johnson-Laird and Byrne (2002) “describe how a conditional's context can affect its interpretation, and ...outline the different possible interpretations of conditionals that these effects of semantic and pragmatic modulation can yield” and did not fail to mention that “In both cases, by definition, we are no longer dealing with basic conditionals.” (p. 658, italics added). When we consider peoples' interpretation of ordinary (non-basic) conditionals, the processes of semantic and pragmatic modulation cannot be bypassed. Their theoretical significance is one of providing explanatory principles that allow us to capture the import(ance) of semantics and pragmatics, content and context. It will be clear that human reasoning, even in highly semantically-impoverished experimental settings, is always performance in situ. When taking pragmatics into account (which in practice, i.e., in the practice psychological research, one cannot not do) one cannot justly say that the interpretation (vs. core meaning) of the conditional is or must be the material conditional, only that is can be the material conditional. Johnson-Laird and Byrne (2002) explicitly noted that their “aim is to make the theory of meaning independent of the theory of comprehension” (p. 648). Evans et al.'s misleading guilt-by-association argumentation is in part due to a failure to recognize this difference between the meaning of a proposition, and the interpretation of a sentence.

Acknowledgements

Our work was done with the support of the Flanders (Belgium) Fund for Scientific Research (W. Schroyens and W. Schaeken) and the Research Council of the University of Leuven (W. Schroyens). We would like to express our appreciation to David Over for his responsibility and willingness to engage in honest and open discussions as regards issues touched upon in the present paper.

BIBLIOGRAPHY


Guilt by Association: On Iffy Propositions and the Proper Treatment of Mental...
NOTES

1. The Duham-Quine thesis in philosophy of science attests to the idea that any prediction is derived from a conglomerate of premises, hypotheses and that, hence, a falsification of the prediction can not be taken to falsify one specific assumption.

2. Evans et al. also state that "it is unclear whether they are fully justified in making exactly this prediction, given the mental footnote element of their theory" (Evans et al., 2003, p. 325; comments added). This reflects undue suspect. They themselves say: "The theory is not determinate at the level of individual participants" (Evans et al., 2003, p. 324; see, Johnson-Laird & Byrne, 2002, p. 654-655). For the purpose of our argument it suffices the conjunction-hypothesis is accepted as being consistent with (vs. predicted by) mental-models theory. Oberauer and Wilhelm (2003) correctly considered that the conjunctive-probability evaluations are predictable within mental-models theory.

ABSTRACTS

We show that by failing to distinguish between logical versus psychological predictions, Evans, Handley, and Over (2003) created strawman arguments that make mental-models theory (MMT; i.e., Johnson-Laird & Byrne, 2002) guilty by its association with assumptions that are known to be psychologically invalid. We also illustrate that MMT allows for different interpretations of ‘if’, including the conjunctive (‘p and q’) or probabilistic (‘if p then possibly q’ or ‘possibly if p then q’) interpretation allegedly beyond the capacity of MMT, and conclude that science is not advanced by a cursory reading of (as such oversimplified) theories that are consequently easy to refute.

On montre que l’absence de distinction entre prédictions logiques et prédictions psychologiques chez Evans, Handley et Over (2003) a donné naissance à de pseudo-arguments qui rendent la théorie des modèles mentaux (TMM) de Johnson-Laird & Byrne (2002) coupable par association avec des hypothèses connues pour être invalides psychologiquement. On montre aussi que la théorie des modèles mentaux permet différentes interprétations du « si », parmi lesquelles l’interprétation conjonctive (‘p et q’) et l’interprétation probabiliste (‘si p, alors peut-être q’ ou ‘peut-être si p, alors q’) qui sont selon Evans et al. hors de portée de la théorie des modèles mentaux. Il apparaît en conclusion que la science n’a rien à gagner d’une lecture rapide et simplificatrice de théories dont la réfutation devient du coup aisée.

INDEX

Keywords: conditionals, mental models theory, interpretation, meaning, if, conditional probability
AUTHORS

WALTER SCHROYENS
University of Leuven, Psychology Department, Lab. Experimental Psychology
Tiensestraat 102,
B-3000 Leuven.
Walter.Schroyens@psy.kuleuven.ac.be

WALTER SCHAeken
University of Leuven
Psychology Department
Lab. Experimental Psychology
Tiensestraat 102,
B-3000 Leuven.
Walter.Schaeken@psy.kuleuven.ac.be