

# A Critical Review of Thinking About What is True, Possible and Irrelevant in Reasoning from or Reasoning About Conditional Propositions

Walter Schroyens (Walter.Schroyens@Ugent.be)

University of Gent.

2, H. Dunantlaan, 9000 GENT, Belgium

## Abstract

Truth-table tasks probe for interpretations of <if A then C> by having people evaluate truth-table cases (e.g., <A\_C>, <A\_not-C>, <not-A\_C>, <not-A\_not-C>) as making a rule true/false, be consistent or irrelevant to the rule. Recent studies and reviews looking at truth-table task performance have used their conclusions about the 'facts' presented in the literature to evaluate the explanatory adequacy of theories of human reasoning. Mental-model theory (e.g., Johnson-laird & Byrne, 2002) has recently been the focus of many critiques. Focusing on those grounded by truth-table task results, our meta-analyses suggest these critiques are in need of a critical analysis as critics have presented mistaken generalizations as fact, constructing thus arguments that are at best unsound.

## General Introduction

Our ability to think conditionally illustrates a quintessentially human marvel; our ability to explain the past and present and to foresee the future by a process of hypothesizing and projecting potential consequences of antecedent events. Reasoning about conditionals has accordingly attracted the interest of cognitive scientists of many disciplines. The list of directly relevant papers literally runs in the thousands.

We aim to provide two contributions to the literature: empirical and theoretical. First, we present a meta-analytic review of literature on the so-called truth-table task, which is one of the main experimental paradigms used to investigate propositional connectives like the conditionals 'if', 'only if', 'if and only if'. It probes for people's interpretation of conditional utterances by asking them to indicate which contingencies are consistent with the conditional or make it true or false. There is a clear need for this review. More than a decade has passed since the last semi-systematic review. Evans and Over's (2004) most recent work that is solely dedicated to 'if' does not provide a new review either and refers to the older work of Evans, Newstead, and Byrne (1993). Moreover, all extent reviews are semi-systematic, that is, they were incomplete and/or conducted by means of at best a counting method, which is known to be in danger of leading to flawed generalizations (Rosenthal, 1983). Our meta-analyses will indeed show that this danger has become a reality and that in recent literature hasty generalisations have been presented as fact.

The second intended contribution is theoretical. The hasty generalizations have been used to argue against one of the most important theories in the field, i.e., the so-called mental-models theory (Johnson-Laird & Byrne, 1991, 2002). Any argument is suspect when based on imprecise generalizations. These arguments are therefore in need of a

critical analyses and revision. "If a successful or at least good argument for a position is subsequently found by any participant to be flawed in a way that raises new doubts about the merits of that position, one is obligated to reopen the issue for further consideration and resolution." This is the 'reconsideration principle' Damer (2005, p. 7) advanced in his work on "Attacking faulty reasoning: A practical guide to fallacy-free arguments". The potential theoretical impact of our meta-analyses goes further and extends to all arguments based on truth-table task results, whether against or in favor of one or the other theory. Evans and Over (2004) stated for instance: "Let us recap some of the experimental evidence that has been presented in this book for the suppositional nature of 'if'. The most direct evidence comes from two main sources: studies of the how people judge the truth value of conditionals in truth table tasks (Chapter 2) and studies which ask people to assess the probability of conditional statements (Chapter 8)". The suppositional nature of 'if' has undoubtedly been the hottest topic in the last five year of research on human reasoning. If the evidence for the suppositional nature of 'if' is so strongly tied to the empirical evidence on the truth-table task, it seems warranted to engage in critical analysis of this task to establish the firmness of the empirical foundations underneath the suppositional theory, which we will discuss in more detail below.

## The truth-table task: A meta-analytic review

Consider a conditional of the form <if A then C>. It expresses a relation between an Antecedent <A> and a Consequent <C>. Both the antecedent and the consequent can be either true or false, implying we have four possible truth contingencies:

TT<A and C>      TF<A and not-C>  
FT<not-A and C>      FF<not-A and not-C>

These truth contingencies are classically referred as TT, TF, FT and FF ('T': Truth or 'F': Falsity of, respectively, the antecedent and consequent.)

In a truth-table evaluation task, people evaluate the four truth-contingencies. In some tasks, people are instructed to reason from conditionals (see Barrouillet, Gauffroy, & Lecas, in press, for a discussion of reasoning from the truth of propositions as opposed to reasoning about the truth of propositions). When reasoning from conditional, the conditional is assumed to be true and participants have to indicate whether the contingencies are consistent or inconsistent with it; that is, they judge whether the contingencies are possible or impossible given that the conditional is true. In other tasks, participants are instructed to reason about the conditional. One does not know if the conditional is true and has to judge whether the

contingencies make <if A then C> true or false, or are irrelevant to its truth-value. Evans and Over (2004) makes a similar distinction between the two evaluation-task formats. We use the theoretically neutral 'two-option' vs. 'three-option' terminology. Three-option tasks include 'irrelevant' as a third response alternative.

## Method

**Literature search.** To locate relevant studies we accessed the PsycINFO, Philosopher's Index and JSTOR data base on the basis of the search terms 'conditional reasoning', 'propositional reasoning', 'sentential reasoning', 'deductive, reasoning', 'implication, reasoning'. The papers were screened for the presence of experiments conducted with the truth-table paradigm and retained upon the condition that they presented mean selection rates for the different response alternatives.

**Table 1.** Response frequencies in different formats of the truth-table tasks with an explicit (E) or Implicit (I) type of reference (cf. reference list for the study)

Study	Study Exp/Condition	Ref.	Options	N	TT		TF		FT		FF	
					T/P	I	T/P	I	T/P	I	T/P	I
1.	'if, then' conditionals	I	2:T/F	50	.83		.03		.05		.55	
2.	Experiment 4	I	2:P/IP	153	.92		.21		.54		.87	
3.	Abstract materials	E	2:C/IC	176	.94		.07		.13		.78	
4.	Experiment 2a.	E	2:P/IP	39	.97		.08		.31		.90	
	Experiment 2b	E	2:P/IP	31	.97		.10		.39		.90	
	Experiment 3	E	2:P/IP	13	.92		.15		.62		.92	
5.	2-option task	E	2:P/IP	100	.91		.19		.57		.87	
	3-options task	E	3:T/F/I	100	.84	.11	.24	.09	.08	.29	.38	.54
	Primed 3-options task	E	3:T/F/I	150	.93	.05	.04	.07	.13	.21	.41	.41
6.	Implicit, "if p then q"	I	3:C/Co/I	40	.96	.03	.06	.04	.05	.33	.29	.63
	Explicit, "if p then q"	E	3:C/Co/I	40	.98	.01	.07	.04	.06	.28	.66	.25
7.	Implicit, "if p then q"	I	3:C/Co/I	25	1.0	.00	.04	.04	0.0	.44	.04	.76
	Explicit, "if p then q"	E	3:C/Co/I	26	.96	.00	.16	.15	.07	.30	.61	.26
8.	Abstract materials	I	3:C/Co/I	120	.98	.00	.01	.06	.06	.53	.06	.88
9.	Temporal order 1	I	3:T/F/I	16	1.0	.00	0.0	.00	0.0	.38	0.0	.88
	Temporal order 2	I	3:T/F/I	16	1.0	.00	.06	.00	0.0	.38	.063	.88
10.	Adults participants	I	3:T/F/I	47	.79	.21	.00	.05	0.0	.82	.10	.90
11.	Past tense	E	C/IC/I	46	.99	.01	.00	.03	.01	.26	.55	.38
	Future tense	E	C/IC/I	46	.99	.00	.01	.00	.03	.22	.57	.37
12.	Factual if	E	C/IC/I	44	.99	.01	.04	.01	.03	.11	.58	.28
13.	Implicit	I	T/F/I	47	.91	.02	.17	.11	0.00	.64	.06	.85
	Explicit	E	T/F/I	48	.75	.06	.15	.10	.06	.54	.15	.58
	Explicit	E	NF/F/I	46	.80	.15	.17	.04	.41	.26	.67	.24
14.	Exp 3	E	P/IP/I	35	.97	.03	.00	.06	.23	.29	.43	.46
	Exp 3	E	T/F/I	31	.77	.13	.03	.16	.13	.39	.29	.42
	Exp 4: +Justification	E	P/IP/I	55	.96	.02	.02	.02	.31	.07	.62	.31
	Exp 4: +Justification	E	T/F/I	51	.96	.02	.00	.04	.10	.33	.47	.45
	Exp 5: from <sup>b</sup>	E	3:P/IP/I	43	1.0	.00	.00	.00	.23	.16	.56	.35
	Exp 5: about	E	3:P/IP/I	39	.97	.03	.03	.08	.21	.26	.44	.38
	Exp 5: from	E	3:T/F/I	38	1.0	.00	.00	.03	.03	.37	.55	.39
	Exp 5: about	E	3:T/F/I	38	.97	.00	.00	.08	.08	.26	.42	.50
	Exp 6: about	E	3:P/IP/I	46	.98	.02	.07	.07	.33	.17	.57	.30
	Exp 6: from	E	3:T/F/I	47	.91	.02	.04	.04	.15	.28	.32	.53
	Exp 6: about	E	3:T/F/I	47	.96	.04	.04	.02	.09	.34	.34	.55
	Exp 6: from	I	3:P/IP/I	50	.92	.04	.00	.10	.14	.26	.24	.46
	Exp 6: about	I	3:P/IP/I	50	.92	.02	.10	.16	.28	.34	.22	.42
	Exp 6: from	I	3:T/F/I	49	.96	.00	.06	.10	.02	.33	.00	.73

**Selection Criteria.** We here consider only adult truth-table evaluation task performance about content-neutral indicative <if A then C> conditionals. That is, non-adult, non-evaluation tasks with non-abstract, non-indicative and/or non-<if A then C> conditionals were not included. The effect of introducing negations in the antecedent or consequent (e.g., "if the number is a 2 then the letter is not a C") also falls beyond the scope of the present article. Every one of these variables (task format, problem, content, age, negation ...) has demonstrated effects. These studies were thus excluded to avoid a potentially biased picture of the bench-mark phenomena.. Studies could obviously only be included when they presented sufficient information about the exact evaluation of all four truth-contingencies. Table 1 presents the basic results of all studies included in the meta-analyses. Space limitations prevent us from being complete

and detail to the studies that did not satisfy the selection criteria (but see, Schroyens, 2007).

## Results

Figure 1 presents mean acceptance rates ('true', 'possible', 'consistent'), as a function of reference type (implicit versus explicit) and task format (two options vs. three options). Reference type refers to the use of implicit or explicit denial to establish the link between the conditional and the denied antecedent and/or consequent in TF, FT and FF. For instance, 'B3' is an implicit FF case for <if A then 2>. Studies that use implicit referencing do not use 'not' to convey falsity/denial; instead they use a specific instance of the contrast-class of the negated object to establish denial (e.g., any letter that is not an A). Explicit FF cases make use of explicit negations (i.e. 'not'; as in 'not A and not-2'). When we are dealing with affirmative conditionals, affirmation is explicit (<A> affirms the antecedent in <if A then 2>). We treated each study as a separate observation and

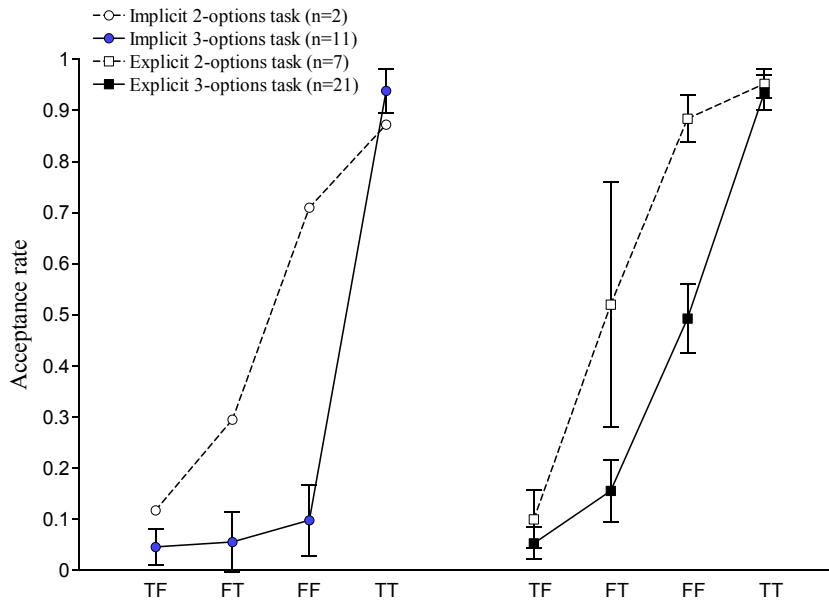


Figure 1: Truth-contingency acceptance rates as a function of reference type (implicit vs. Explicit) and task format (two vs. three-option format).

conducted analyses of variance with the task format and type of reference as between-group variables and the truth contingency as a repeated measurement.

The combined results reveal the following pattern:  $1 \approx P(TT) > P(FF) > P(FT) > P(TF) \approx 0$ . The linear contrast is highly reliable,  $F(1,37) = 1021.01$ ,  $Mse = .0008$ ,  $p < .0000001$ . Table 1 shows every single study conforms to this pattern. Acceptance rates nonetheless shift considerably as a function of task format and/or reference type. Figure 1 clearly reveals two main effects: people are more likely to accept a truth-contingency in the two-option versus three-option task (.347 vs. .556;  $F(1,37) = 46.353$ ,  $Mse = .019$ ,  $p < .0000001$ ), and are more likely to accept a truth-contingency when the reference is explicit (.392 vs. .512;  $F(1,37) = 15.193$ ,  $Mse = .019$ ,  $p < .001$ ). Both the two-option and three-option tasks yields implicitness effects, (respectively: .499 vs. .614,  $F(1,37) = 4.256$ ,  $Mse = .019$ ,  $p < .05$  and .284 vs. .409,  $F(1,37) = 23.163$ ,  $Mse = .019$ ,  $p < .0001$ ) without interacting significantly ( $F < 1.0$ ).

It is clear from Figure 1 that the implicitness effect and the task effect are only observed on false-antecedent cases (FT and FF). Both reference type and task format interact reliably with the truth status of the antecedent; respectively  $F(1,37) = 15.476$ ,  $Mse = .0141$ ,  $p < .001$  and  $F(1,37) = 53.379$ ,  $Mse = .0141$ ,  $p < .0001$ ). That is, combined over FF and FT we have a respectable task-format effect ( $F(1,37) = 53.3703$ ,  $Mse = .031$ ,  $p < .0001$ ) and a sizable implicitness effect ( $F(1,37) = 16.614$ ,  $Mse = .031$ ,  $p < .001$ ). These effects are generally not obtained with true-antecedent cases ( $F$ 's  $< 1.1$ ).

Three-option tasks present 'irrelevant' as a third response option along with 'true\*' and 'false' (we mark true with '\*' to indicate that 'true' is not always the actual response alternative in the three-option task, cf. Table 1). By looking at only acceptance rates (possible vs. true\*), we do not

know whether the task-format effect is reflected in a shift towards 'false' or 'irrelevant' judgements of false-antecedent cases that are deemed 'possible' in the two-option task. We are similarly still ignorant about the exact nature of implicitness effects in the three-option task: People are less likely to accept implicit false-antecedent cases. Does this mean that they make more irrelevancy judgements and/or that they give more 'false' responses? Figure 2 presents the selection rates for all response options in three-option tasks.

First consider the task-format effects separately for explicit or implicit false-antecedent (FT and FF) cases. We have only two studies using two-option tasks with an implicit reference type. That is, we have a less strong empirical basis to generalize the implicitness effect, and therefore only discuss the more reliable response rates on explicit two-option vs. three-option tasks. The TF(Possible) and FF(Possible) rates are .520[.281,.950] and .884[.838,.929] (in brackets we give the 95% reliability interval).

The task-format effect on explicit FF cases, .884[.838,.939] vs. .493[.425,.560] ( $d = .391$ ) is almost completely explained by the .389[.229,.439] selection rate of FT(Irrelevant); the comparable FT(Impossible) and FF(False) rates remain stable across the two explicit tasks; .116[.070,.162] vs. .118[.087,.149]. The picture seems somewhat more complicated when we look at FT. The task-format effect on explicit FT, .520[.281,.950] vs. .155[.094,.217] ( $d = .364$ ), is mostly explained by the FT(Irrelevant) selection rate of .271[.225,.316]. However, part of the task effect seems captured by a slight increase in FT(Impossible) versus FT(False) judgements: .480[.241,.719] vs. .573[.508,.639] ( $d = .097$ ).

Figure 2 shows that the .494[.425,.560] vs. .098[.028,.168] implicitness effect on FF ( $d = .396$ ) is mirrored in the 389[.339,.439] vs. .737[.623,.851] ( $d = .348$ ) increase of irrelevancy judgments about implicit FF cases. The implicitness-effects on FT (.156[.094,.216] vs. .055[.000,.113];  $d = .101$ ) reflects a similar shift from irrelevant judgments for implicit cases towards true judgments for explicit cases (.271[.225,.316] vs. .434[.323,.546]).% of participants answering respectively the first and second compatibility question. We take the complete failure of any subject to successfully apply Handley et al.'s (2006) definition as a suggestion that modal reasoners do not reason in line with this notion of compatibility.

## Discussion

Our meta-analyses establish bench-mark phenomena regarding truth-table evaluation task performance. The overall results pattern shows that most people accept TT, whereas they almost unanimously reject TF. The evaluation of false-antecedent cases is much more mouldable, though invariably FT is rejected more frequently than FF. FT and

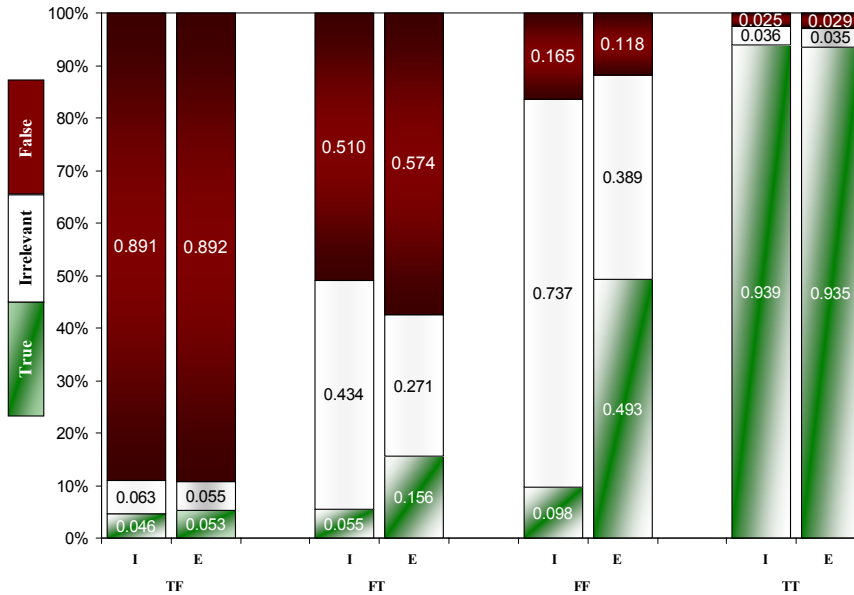


Figure 2: Mean proportions ‘True’, ‘Irrelevant’, and ‘False’ judgements as a function of the type of referencing (implicit vs. Explicit).

FF evaluations are affected both by task format and reference type: The elimination of ‘irrelevant’ in two-option tasks is mostly reflected in an increase of ‘possible’ judgements of false-antecedent cases; the use of an implicit reference causes an increase of ‘irrelevancy’ judgements of these false-antecedent cases. In the following we illustrate how mental-models theory deals with these bench-mark phenomena.

Mental-models theory currently explains irrelevancy judgments on the basis of the so-called implicit-model or initial model principle: people will initially represent only contingencies expressed explicitly: <if A then C> speaks about <A> in relation to <C>. Thus, initially, people only consider <A> in relation to <C>. There is a mental place holder <...> for other possibilities but these contingencies are not represented explicitly, at least not to start with. It has the benefit of minimizing representational costs in a system with limited processing resources. Since the initial-model representation of <if A then C> does not capture FF<not-A\_not-C> or FT<not-A\_not-C>, these contingencies will initially be judged ‘irrelevant’: “There is no explicit mental model representing the possibilities in which the antecedent is false, and so naive individuals and theorists deem the conditional neither true nor false but irrelevant” (Johnson-Laird & Byrne, 2002, p. 658, italics added).

The initial-model principle is consistent with the findings, but is in itself insufficient to explain the variation in irrelevant ratings. For instance, it does not distinguish the two types of ‘irrelevant’ responses that are evidenced by the implicitness effect. The initial representation of <if A then C> would capture the TT<A\_C> contingency, independent of whether people evaluate the explicit FF<not-A\_not-C> contingency or a specific implicit FF case. If we consider what is means for something to be ‘relevant’, we see that both the extra irrelevant evaluations caused by implicit referencing and the basic irrelevant evaluations are readily

explained in mental-models theory. Take “if the figure is a circle, then it is colored red”. The topic of discourse in “a green square” mismatches both the antecedent and consequent. That is, upon a fast and frugal treatment of “a green square” it is unrelated-irrelevant regarding what is talked about in the conditional utterance.

“A figure that is not a circle and that is not colored red” would not be judged unrelated-irrelevant. The topics of discourse in this explicit FF case match the topics of discourse in the conditional (they are both about circles and being colored red). That is, it is not the case that they have nothing to do with one another: There is some link and thus some type of relevance between to two propositions. But, given that we know that there is some link/relevance we still do not know what that relevance actually is. In the absence of an explicit representation of the false-antecedent cases in the initial representation of the

conditional, these are not judged unrelated-irrelevant but unidentified-irrelevant. From the initial representation it is not immediately apparent what the link/relevance would be, and people take a fast and easy way out by saying it is irrelevant. Below we come back to issue of irrelevant evaluations of explicit false-antecedent cases and provide yet another reason why people would judge false-antecedent cases to be irrelevant. The difference between irrelevant evaluations of explicit versus implicit cases demonstrates that mental-models theory initial-model principle is by itself insufficient to account for irrelevant evaluations. There are different degrees of ‘(ir)relevancy’. Something can be completely unrelated, or it can be related but the relevancy of the related cases is not transparent from the initial-representation of the conditional.

In summary, mental-models theory can explain the major pattern of results in the truth-table tasks. The initial-model principle, including the implications of mental-footnotes about what is false, explains the relatively high irrelevant ratings of the false-antecedent cases, FF<not-A\_not-C> and FT<not-A\_C>; the high false rates of TF<A\_not-C> under both a conditional or biconditional interpretation of the conditional, as well as the non-negligible proportion of false evaluations of FT<not-A\_C> under a bi-conditional interpretation. When the link between the truth-table cases is implicit, there is mismatch between the topics of discourse in these contingencies on the one hand, and the topics of discourse in the conditional on the other hand. The implicit (i.e., mismatching) truth-table cases are consequently often evaluated as unrelated/irrelevant. This is the robust implicitness effect shown in Figure 2. Johnson-Laird and Byrne’s (2002) explicit discussion of differences between reasoning about truth versus reasoning about possibilities shows that mental-models theory is consistent with differences between the possibility-based and truth-based truth-table tasks. Most interestingly, the distinction between

reasoning from and reasoning about propositions gives rise to further potential processing difficulties in reasoning about the truth of propositions. Below we present further analyses of truth-based tasks, which confirm this novel hypothesis and establishes a new phenomenon.

## General Discussion

Others also looked at truth-table task performance, and used their own conclusions about the 'facts' presented in the literature to evaluate the explanatory adequacy of theories of human reasoning. Mental-models theory (e.g., Johnson-laird & Byrne, 2002) has recently been the focus of many critiques. In looking at critiques grounded by truth-table task results it becomes apparent that there is need for critical re-analysis. Our meta-analyses show that some critics presented imprecise generalizations as fact and thus constructed arguments that are at best unsound (i.e. valid but with false premises).

The issue at stake is what we can call the irrelevancy effect. The irrelevancy effect refers to the observation that when given the opportunity, people judge false-antecedent cases,  $FF\langle \text{not-A}_C \rangle$  and  $FF\langle \text{not-A}_\text{not-C} \rangle$  as irrelevant regarding the truth of  $\langle \text{if A then C} \rangle$ . In fact, the crucial datum providing the empirical basis of 'crucial' arguments against mental-models theory is even more specific. The issue at stake is how many people will make such an irrelevancy judgement of especially the  $FF\langle \text{not-A}_\text{not-C} \rangle$  contingency. Our meta-analyses show that the empirical claims in Evans et al. (2005, p. 1043), Evans and Over (2004, p. 76, p. 239), Evans et al. (2003), Handley, Evans, and Thompson (2006, p. 564), and Evans (2007, p. 386) are wrong and demonstrate that the authors of these papers made a hasty generalization in presenting the 'facts' presented in the literature in such a way as to suggest a majority of people think  $FF\langle \text{not-A}_\text{not-C} \rangle$  cases are irrelevant regarding the conditional.

It may be noted that Evans and Over (2004) do recognize that not all people will demonstrate a so-called defective truth-table in which both  $FT\langle \text{not-A}_C \rangle$  and  $FF\langle \text{not-A}_\text{not-C} \rangle$  cases are irrelevant while  $TT\langle A_C \rangle$  is judged 'True' and  $TF$  is judged 'false'. The explanatory adequacy of Evans et al's own theory is however quite irrelevant when we are analysing claims about (a) the empirical observations, i.e., 'facts' about  $FF\langle \text{not-A}_\text{not-C} \rangle$  evaluation rates and (b) the negative theoretical implications they argued exist for alternative theories in virtue of these facts. We do not engage in c) a critique of Evans and Over's (2004) own theory and the proclaimed positive implications the empirical observations would have for this theory. Space does not allow us to tackle (c); instead we aim for a thorough treatment of (a) and (b). First, our meta-analyses show that Evans et al. were mistaken about the facts. Second, given that they got the facts wrong we have the intellectual obligation to reconsider arguments based on these mistaken facts.

The idea that a majority considers  $FF\langle \text{not-A}_\text{not-C} \rangle$  irrelevant has been a central premise in a recurring argument Evans et al. (2005; Handley et al., 2006; Evans, 2007) make against mental-models theory. Let us consider the argument in somewhat more detail.

We find this [model theory] account [of the irrelevancy effect] highly implausible for the following reasons. When presented with a minor premise for the MT argument, people must succeed in fleshing-out the not-p and not-q cases as a true possibility around 70% of the time to explain observed success rates in the experiments. Recall our example: If there is an A on the card then there is a 3 on the card. There is not a 3 on the card. How is it then possible that, in the truth table task, a not-A and not-3 case is evaluated as irrelevant on the majority of occasions?" Evans et al., 2005 (p. 1043, italics added).

This cross-task argument hinges on the premise that according to Evans et al.'s interpretation of the mental-models theory it should predicted a match between the acceptance rates of  $FF\langle \text{not-A}_\text{not-C} \rangle$  and acceptance rates of  $MT\langle \text{not-C therefore not A} \rangle$

Evans et al. (2005) state MT "is made ...perhaps about 60-70% of the time". Proper meta-analyses yield a more accurate estimate of .699[.671;.728] (Schroyens et al., 2001a, 2001b, Schroyens & Schaeken, 2003). Hence, according to Evans et al's (2005) cross-task argument, people would need to judge  $FF\langle \text{not-A}_\text{not-C} \rangle$  as a true possibility about 70% of the time. Evans et al. (2005) state 'irrelevant' is the dominant evaluation. General readers would therefore accept the invited inference that the cross-task prediction fails, and that the model theory is therefore fundamentally flawed or faced with a "serious problem". With a presumed majority of  $FF\langle \text{Irrelevant} \rangle$  evaluations, we would have considerably less than 50% of people accepting  $FF$ . Hence, the model theory would be faced with an unexplained discrepancy with a lower-limit around 20% (.699 vs. minimally .50).

We can reconstruct the cross-task argument as a valid argument. It is based on two premises. There is the theoretical premise that if the model theory is right, then  $FF$  and  $MT$  rates should match and there is the empirical premise signifying denial of the theoretical premise's consequent (i.e.,  $FF$  and  $MT$  rates do not match). It follows (logically, via  $MT$ ) that the model theory is wrong. Though valid, it is not a sound argument (i.e. valid with true premises).

The above quote shows that Evans et al. (2005) use an example with explicit negations to illustrate their argument ('not-3' denies '3' explicitly). Of course we cannot be comparing proverbial apples and oranges: We need to consider explicit cases in both the truth-table task and the inference task. When doing so, the data clearly show that explicit  $FF$  is not typically evaluated as 'irrelevant';  $FF\langle \text{Irrelevant} \rangle = .389[.339, .439]$ . Evans et al.'s (2005) imprecise (over)generalization has the implication that all arguments based on this false assumption are suspect. They are potentially rendered null and void; thus the need for critical revision (cf. the abovementioned reconsideration principle: Damer, 2005).

A thorough treatment of the literature shows that one can only claim irrelevancy as a dominant response when limiting the focus to implicit truth-table studies. This obviously also implies -- as again we should not be comparing apples and oranges -- that if Evans et al. want to

make their argument, they need to consider the results on implicit MT arguments. (Note this limited scope already weakens the cross-task argument). As far as we know, not a single study exists that makes MT fully implicit, analogous to the fully implicit nature of ‘B3’ vis-à-vis “if A then 2”. That is, a non-confounded comparison of the implicit truth-table and implicit inference task cannot be made. A fully implicit MT argument would go: <if A then 2; 3, therefore B>. It would almost be folly (i.e. too trivial) to run such a study. Were we to run the study, then MT acceptance rates would undoubtedly be much lower than .699. Indeed, the fact that the first author is not American (not-A) does in no way entail that he is Belgian (B). Using “B” or any other specific contrast-class element of ‘not-A’ as the conclusion transforms MT into a logically invalid argument. In short, not a single matched inference task exists with which to compare the dominant ‘irrelevancy’ response about implicit false-antecedent (i.e. FF) cases.

If the cross-task argument can not be applied to implicit cases, maybe it can still be applied to explicit cases. The basics of experimentation teach us that we control for potential confounds. That is, explicit truth-contingency evaluations are compared with explicit inference ratings. When we look at the actual data, we see that the explicit FF contingency is judged True\* in about .493[.425,.560] of cases and that it is considered ‘possible’ in about .883[.838,.930] of cases. It is not entirely clear whether ‘true\*’ or ‘possible’ is the most appropriate predictor for the .699[.671;.728] explicit MT rates.

We must here satisfice with the negative conclusion that Evans et al (2003, 2005) argument is rendered suspect (since based on a false premise), and cannot engage in a more constructive critical analysis of the cross-task argument. Supposing for arguments sake that the cross-task argument is valid, meta-analyses show it is unsound. But, is the cross-task argument really a valid argument? That is, do we really need to expect that FF<not-A\_not-C> evaluations match MT<not-C therefore not-A> inference rates? A critical analysis of mental-models theory indicates that at least within this theory FF<not-A\_not-C> and MT<not-C therefore not-A> evaluations do not need to match. At best one can expect a positive relation.

## References

- <sup>10</sup>Barrouillet, P., Gauffroy, C., & Lecas, J.-F. (2006). Is the mental model theory of conditional really mistaken? Reasoning from, and reasoning about conditionals by mental models. Paper submitted for publication, Université de Bourgogne, France.
- <sup>11</sup>Byrne, R. M. J. & Egan, S. M. (2004). Counterfactual and prefactual conditionals. *Canadian Journal of Experimental Psychology*, 58(2), 113-120.
- Damer, E. T. (2005). *Attacking faulty reasoning: a practical guide to fallacy-free arguments*. (5th ed). Belmont, CA: Thomson Wadsworth.
- <sup>12</sup>Egan, S. M., Byrne, R. M. J., Garcia-Madruga, J. A. (2007). Factual and counterfactual ‘only if’ conditionals. Unpublished paper submitted for publication.
- <sup>6</sup>Evans, J. St. B. T. (1983). Linguistic determinants of bias in conditional reasoning. *Quarterly Journal of Experimental Psychology*, 35A, 635-644.
- <sup>7</sup>Evans, J. St. B. T., Clibbens, J., & Rood, B. (1996). The role of implicit and explicit negation in conditional reasoning bias. *Journal of Memory and Language*, 35, 392-409.
- <sup>8</sup>Evans, J. St. B. T., Handley, S. J., Niens, H., Over, D. (2006). The mental representation of abstract conditional sentences: A study of qualitative differences. Unpublished paper, University of Plymouth, Plymouth, UK.
- Evans, J. St. B. T., Handley, S. J., & Over, D. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29(2), 321-335.
- <sup>9</sup>Evans, J. St. B. T., & Newstead, S. E. (1977). Language and reasoning: A study of temporal factors. *Cognition*, 5, 265-283.
- Evans, J. St. B. T., & Over, D. E. (2004). *If*. New York: Oxford University Press.
- Evans, J. St. B. T., Over, D., & Handley, S. J. (2005). Suppositions, Extensionality, and Conditionals: A critique of the mental model theory of Johnson-Laird & Byrne (2002). *Psychological Review*, 12(4), 1040-1052.
- <sup>2</sup>Handley, S. J., Evans, J. St. B. T., & Thompson, V. A. (2006). The negated conditional: A litmus test for the suppositional conditional. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32 (2), 559-569.
- Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109(4), 646-678.
- <sup>1</sup>Paris, S. G. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology*, 16, 278-291.
- <sup>13</sup>Schroyens (2007). The folly of presuming the induction problem solved: A critical analysis of the truth-table paradigm. Paper submitted for publication, University of Gent, Gent, Belgium.
- <sup>4</sup>Schroyens, W., Handley, S., Evans, J. St. B. T., & Schaeken, W. (2002). Evaluating conditional arguments: Opposites do not attract. (Report No. 292). Leuven: University of Leuven, Laboratory of Experimental Psychology.
- <sup>5</sup>Schroyens, W., & Sevenants, A. (2006). Issues in Reasoning about Iffy Propositions: Reasoning times in the Syntactic-Semantic Counter-Example Prompted Probabilistic Thinking and Reasoning Engine (SSCEPTRE). In Proceedings of the 28th Annual Conference of the Cognitive Science Society New Jersey: Lawrence Erlbaum.
- <sup>14</sup>Sevenants, A. (2007). Reasoning from and reasoning about conditionals. Unpublished data, Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium.
- <sup>3</sup>Thompson & Byrne, 2002. Reasoning counterfactually: Making inferences about things that didn’t happen. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 28(6), 1154-1170.