The Strategy Behind Belief Revision: A Matter of Judging Probability or the Use of Mental Models?

Ann G. Wolf (ann.g.wolf@psychol.uni-giessen.de)
Department of Experimental Psychology, Otto-Behaghel-Straße 10F
D-35394 Gießen, Germany

Markus Knauff (markus.knauff@psychol.uni-giessen.de)
Department of Experimental Psychology, Otto-Behaghel-Straße 10F
D-35394 Gießen, Germany

Abstract

Research in the field of human reasoning has shown repeatedly that people find it reasonably easy to detect inconsistencies. The question that still remains is how people revise their beliefs to undo these inconsistencies. We report two experiments in which subjects had to make belief revision choices on modus ponens (MP) and modus tollens (MT) sets of problems that contained conditionals with different levels of probability. After the final statement of each set, which was stated to be true, they had to decide which of the first two statements they believed more. The results showed that with high and low probability problems, subjects revised their beliefs as a function of the probability of the conditional. However, when the conditional had a near 50% chance of occurring, the pattern of belief revision more resembled the mismatch principle as suggested by the mental model theory. Thus, people use different strategies when revising their beliefs depending on whether they are guided by semantics or mental models.

Keywords: Belief revision, conditional probability, mental model theory, mismatch principle.

Introduction

Belief revision is the process by which one alters his or her belief state in the face of contradicting evidence. In the past, the topic of belief revision was reserved for logicians and philosophers to be concerned about (e.g. Gärdenfors, 1988). Several opposing theories have been offered in the past to explain the underlying process of belief revision. The most long-standing theory of belief revision is that of epistemic entrenchment, developed within the field of artificial intelligence (e.g. Gärdenfors, 1992). This notion holds that some pieces of information are more entrenched and are accordingly more easily believed than others. There have been some findings suggesting that epistemic entrenchment plays a role in belief revision. However, it is not clear exactly what defines epistemic entrenchment. What the current study aims to investigate is whether probability of conditionals serves as a kind of epistemic entrenchment in belief revision and to contrast this with the mental model theory of belief revision (Johnson-Laird, 2006; Johnson-Laird, Girotto, and Legrenzi, 2004). Elio and Pelletier (1997) were one of the first to incorporate the theory of belief revision into human deductive reasoning. They explored conditional reasoning wherein Modus Ponens (MP) is the logic inference rule of the sort ‘if p then q’ where conclusion q is inferred from categorical statement p. Similarly, Modus Tollens (MT) is the logic inference rule of the sort ‘if p then q’ where conclusion not-p is inferred from the categorical statement not-q. They found that for both MP and MT inferences, the preferred change was disbelieving the conditional statement. They ascribed this finding to syntactical factors. Politzer and Carles (2001) investigated whether the conditional form of the first premise in MP and MT problems influences the relative entrenchment of the conditional and categorical premises by also using disjunctive and conjunctive syllogisms. In line with Elio and Pelletier’s (1997) findings, the conditional premise was revised more often than the categorical premise. Because the disjunctive and conjunctive major premises were also more questioned than their minor counterpart, the authors argue that is not the conditional nature of the major premise leading to these revision choices but its status quo of being the major premise, by which they mean that due to its compound nature its more likely to be the source of error. Elio (1997; second experiment) also found that the conditional statement was believed less than the categorical. The accounts put forward by Elio and Pelletier (1997) and Politzer and Carles (2001) for explaining the deeper entrenchment of the conditional are context independent (Hasson & Johnson-Laird, 2003). A point to remark here is that although they used non-basic inference problems and the accounts offered seem plausible within the context of their studies, the thematic content of their premises were science fictional. In real life, however, people reason about more mundane issues. Byrne and Walsh (2005) argued that relative entrenchment is a function of familiarity. They found that with problems in an unfamiliar domain, people tended to believe the categorical premise more. In contrast, with familiar problems people tended to believe the conditional more. The reason for this, they say, is that familiar conditionals may express rules and laws that are deeply entrenched and therefore people rather choose to revise the categorical fact. However, the familiar conditionals they used had a high intrinsic probability close to being factual of nature, therefore stripped from any subjective probability. Not all familiar statements express rules and laws and a familiar conditional need not necessarily be high in probability. With respect to the
unfamiliar domain, people might find it difficult to assign any level of probability to the conditionals because they cannot imagine the situation sketched before them.

Despite these findings and explanations, what is still missing is a firm theoretical basis for human belief revision performance in conditional reasoning. The first to offer such a theory were Johnson-Laird and Byrne (2002) which will be outlined in the next section.

The Mental Model Theory of Belief Revision

According to the mental model theory of reasoning, people construct a set of mental models of the possibilities that the situation embedded in the premises might represent. The key assumption of the mental model theory is that mental models represent only what is true according to the premises, and not what is false, which is called ‘The principle of truth’ (Johnson-Laird, et al., 2004). Initially, people construct only one possible mental model, the conjunctive p & q, which is called the explicit model. They do however make ‘mental footnotes’ of further implicit models that if fully fleshed out represent the remaining false possibilities, which are “not-p and q” and “not-p and not-q”. Johnson-Laird and his colleagues posit that the first step in reasoning is detecting an inconsistency among a set of premises (Johnson-Laird et al., 2004). They put forward the principle of models of consistency which holds that people search for a mental model that holds a possibility in which all premises are true (Legrenzi, Girotto, & Johnson-Laird, 2003). If they find such a model then the set of premises is judged as consistent, otherwise it is regarded inconsistent.

How does the mental model explain how people resolve inconsistencies? The model theory uses a so-called mismatch principle to explain and predict which of the two statements will be revised. According to this principle, the statement that will be revised or believed less is that statement, whether it be conditional or categorical, that has a mental model that not only mismatches but also conflicts with the mental model of the contradicting fact (Hasson & Johnson-Laird, 2003; Johnson-Laird, 2006; Johnson-Laird, et al., 2004). With MP problems, the contradicting fact would be not-q, this mismatches and conflicts with the mental model of the conditional (p q) and is therefore discarded. With MT problems, the contradicting fact is p, which matches the model of the conditional and therefore people would revise the categorical instead. Johnson-Laird and his colleagues (Johnson-Laird, 2006; Johnson-Laird et al., 2004) demonstrate the strength of their theory by harmonizing the results of former studies with the mismatch principle. For example, Elio and Pelletter (1997) found that the belief revision was a function of which counter fact followed the belief set. When it was a negation of the consequent, then subjects tended to reject the conditional and believe the categorical statement more. However, when the counter fact was of the form p then they believed the conditional statement more.

The Probability of Conditionals

A number of authors have claimed that people represent ordinary conditionals psychologically (e.g. George, 1995; Liu, Lo, & Wu, 1996; Oaksford & Chater, 2001; Stevenson & Over, 1995). People interpret a conditional “if p then q” as ‘what is the probability of q given p’ and this seems to affect their reasoning strategy. For example, in a study by Liu, et al. (1996), subjects had to decide whether the conclusion followed logically from the premises on valid and invalid inference problems. They found that the higher the perceived sufficiency of the problems, the higher the correct responses. In a follow-up test, they also found a positive relationship between the probability judgment of the conditional and the endorsement of the set of premises. Evans, Handley, & Over (2003) have shown in a similar fashion that people are less likely to endorse a conditional when the antecedent has a low probability.

The influence of probability of conditionals on reasoning came to be known as the conditional probability hypothesis, first implied by Marcus and Rips (1977) and later further developed by other researchers (e.g. Over, Hadjichristidis, Evans, Handley, & Sloman, 2007). The conditional probability hypothesis is grounded on the conditional subjective probability, P (q|p), which is known as the Ramsey test (Edgington, 1995, as cited in Oberauer & Wilhelm, 2003). This test implies that people judge the conditional the same as the conditional probability. Several authors have contrasted the predictions of the mental model theory with those of the conditional probability hypothesis on reasoning tasks. The findings in these studies supported the conditional probability hypothesis and not the material conditional hypothesis. This was the case in studies using basic conditionals (see Johnson-Laird & Byrne, 2002) with specified frequency distributions (Evans, et al., 2003; Oberauer & Wilhelm, 2003) as well as in studies using non-basic conditionals (Over & Evans, 2003; Over et al., 2007).

The Current Study

What we intend to do here is to extend the conditional probability hypothesis into the research on belief revision and test it against the mental model theory of belief revision. We propose here that if people judge conditionals in a probabilistic manner, they also should take the probability of the conditional into account in their process of belief revision and base their reasoning thereon. We will use sets of conditional problems statements that relate to real-life situations with differing levels of probability.

We hypothesize that when the “if p then q” relationship in the conditional has a clear high or low chance of occurring then people will use this information as a sort of heuristic to guide their belief revision. When the premises give off no sign of which one would be more probable and thus more believable, people will turn to the use of mental models to resolve the inconsistency as they would with basic conditionals. Two experiments were carried out to test the mismatch principle and the conditional probability hypothesis in belief revision. The first experiment
investigated whether people use subjective probability as a means to guide their revision choice. To test this, we used inference problems with either high or low probability conditionals. We expect that people will believe the conditional more when it expresses a high probability and the categorical more when it expresses a low probability. Further, when the probability of the conditional is close to 50%, then people will convert to the use of mental models to guide their revision choice.

**Experiment 1**

**Methods**

**Subjects.** Thirty students (17 females, 13 males) aged 19 to 31 from the University of Giessen participated in this study in exchange for monetary compensation.

**Materials and Design.** The conditionals used in the experiment were taken from a pool of 36 conditional statements. These conditional statements were rated for probability by another group of students that were taken from the same population. These students were recruited in seminars and tutorials of a first year course and were tested in the class rooms. They were given a booklet that consisted of 36 individual conditional “If-then” statements, each presented on a separate page. On the first page of the booklet, they read the following instruction (translated from German):

“On each of the following pages you will be presented with a statement, which is uttered by a person. A rating scale is presented under each statement. On this scale, please rate the possibility that this person is speaking the truth. 0% means “very unlikely” and 100% means “very likely”. Please rate the statements in the order they are being presented to you. Please do not go back to previous statements.”

Following Girotto and Johnson-Laird’s (2004) suggestion, we asked the participants how possible it is that the statement is true. The statements were randomly presented to the subjects and they were allowed to work at their own speed.

Descriptive statistics were run over all the statements to find the eight statements with the lowest mean and the eight statements with the highest mean. The overall mean of the low-probability statements was 7.19 and the overall mean of the high-probability statements was 88.44. A Wilcoxon Signed Ranks Test showed that the difference between the two means was highly significant, $z = -11.030, p < .000$. From each of these 16 statements, a set of MP problems and a set of MT problems were generated resulting in 32 problem sets in total. Thus, we used a 2 x 2 design with Inference problem (MP vs. MT) and Probability of the conditional (low vs. high) as the independent within subject factors. The dependent factors were revision choice and decision time. The presentation software package Super lab 4.0 was used for presenting the items on the computer screen. For the MP set of premises, the third sentence was always a negation and for the MT set of premises, the third sentence was always an affirmation.

Example of a high probability MT:  
If Christian is fishing, then he is quiet. (p q)

Example of a low probability MP:  
If Knut goes to work, then he will take a hot air balloon. (p q)

Knut goes to work. (p)

Christian is not quiet. (¬q)

He does not take a hot air balloon. (¬q)

Christian is fishing. (p)

The presentation of the 32 problem sets was randomized across the subjects.

**Procedure.** Subjects received instructions on the computer screen. They were explained that they would be presented with three statements one at a time and that the truth status of the first two statements was uncertain but that the third statement, which was inconsistent with the first two, was certainly true. Their task was then to decide which of the first two statements they believed more. Four practice trials preceded the actual experiment. They had the possibility to take a short break after each set of premises. Both the conditional and the categorical statement appeared in white font colour in the middle of a black screen. After reading each statement, the subjects had to push a spacebar-like button to continue. Next, the inconsistent statement that was known to be true appeared on the screen; the fact was printed in red font to signify its counterfactual nature and to aware the subjects they had to make a belief revision choice. For that, they had to press one of two buttons, depending which of the two previous statements they believed more, the conditional or the categorical. The designation of the left and right button to conditional and categorical statements was counterbalanced across subjects.

**Results and Discussion**

Mean percentages of belief revision choices in the four conditions are depicted in Figure 1.

![Figure 1: Belief revision choices in percentages. Notes: MPH = Modus Ponens high-probability; MPL = Modus Ponens low-probability; MTH = Modus Tollens high-probability; MTL= Modus Tollens low-probability.](image)

We used paired-samples t-tests to analyse the impact of probability on belief revision. The conditional statement was believed significantly more often when the MP
statement had a high probability ($M = 61.67\%$, $SD = 30.07$) than when it had a low probability ($M = 22.08\%$, $SD = 21.45$), $t(29)$, $p < .000$. Similar results were found for the MT problems. With the high-probability MT problems, the conditional statement was believed significantly more often ($M = 82.50\%$, $SD = 20.13$) than with the low-probability MT problems ($M = 26.25\%$, $SD = 24.86$), $t(29)$, $p < .000$. These findings show that people take the probability of the conditional into account when they have to make a belief revision choice. This is supportive of the subjective probability hypothesis. The results also extend Politzer and Carles’ (2001) findings by showing that it also holds for MT inference problems. Figure 2 depicts the mean RTs for the 4 conditions.

![Figure 2: Mean reaction times for the 4 conditions](image)

A repeated measures ANOVA was performed to examine the impact of Inference problem and Probability on RT. No main effect was found for Inference problem, $F(1, 29) = .266$, $p = .610$. A main effect did show up for the Probability, Wilks’ Lambda$=.38$, $F = 46.760$, $p < .000$. The decision time was higher for MT problems ($6.16$ s) than for MP problems ($4.73$ s). However, a significant Inference problem x Probability interaction, $F (86.489)$, $p < .000$, accounted for the main effect for Probability. Post-hoc tests revealed that the difference in reaction times between high- and low-probability problems was only significant for the MT inference problems. With low-probability MT problems, it took subjects significantly longer ($M = 7.07$ s) to decide which statement to believe more than with the high-probability problems ($M = 3.95$ s). A possible explanation for this could be that the strategy of finding inconsistencies between mental models might have interacted with the strategy of making probability judgments. Recall that the mismatch principle takes the stand that people will choose to disbelieve the statement that has a mental model that is inconsistent with that of the incontrovertible fact. With both the high- and low-probability MT set of statements, the mental model of the incontrovertible fact (p) matches the mental model of the conditional statement (p and q). But only with the high-probability problems, the subjects believed the conditional more. Here, the strategies predicted by the mental model theory of belief revision and the conditional probability hypothesis work in tandem. This might explain the fast reaction times with the high-probability MT problems; it was most likely the easiest to perform for the subjects. With the low-probability MT problems, subjects chose to believe the categorical statement more than the conditional, which is in accordance with the conditional probability hypothesis. However, the mental model of the given fact (p) does not fit the mental model of the categorical statement (not-q). This conflict could account for the high reaction time with the low probability MT problems.

To test whether probability-based reasoning and the mismatch principle represent two different strategies, we conducted a second experiment with inference problems taken from the first experiment, intermixed with MP and MT inference problems that had a mean probability score close to 50% in the rating study. If indeed the two strategies are both in play and the high versus low probability of the conditionals overruled the strategy according to the mismatch principle, then subject should resolve the inconsistencies on these so-called ‘neutral’ probability problems by revising their belief with the help of mental models. This means that with MP neutral-probability problems, the subjects should choose to believe the categorical statement more because the mental model of the fact (not-q) conflicts with the mental model of the conditional statement (p q). Contrary, with MT neutral-probability problems, subjects are predicted to believe the conditional statement more because its mental model matches the one of the presented fact (p). The belief revision choices with the low- and high-probability problems should show the same trend as in experiment 1.

### Experiment 2

**Subjects.** A total of 40 subjects will be tested. For the completion of this article 16 people aged 18 to 26 were tested (7 men, 9 women) from the University of Giessen in exchange for monetary compensation.

**Materials and design.** The high- and low-probability problem sets were taken from the set used in the first study. Conditionals that were rated around 50% probability in the rating study were used to create the sets of ‘neutral’ problems. A total of 24 sets of problems were used which were divided into the following six conditions: four high- and four low-probability MP problems, four high- and four low-probability MT problems, four ‘neutral’ MP problems, and four ‘neutral’ MT problems. Again, the software program *Super Lab 4.0* was used for presenting the items.

**Procedure.** The instructions on the computer were the same as in experiment 1. The items were presented in a randomized fashion and preceded by four practice items.

### Results and Discussion

The results for the low- and high-probability problems showed a pattern similar to that of the first experiment (see Figure 3). With the MP high-probability problems, subjects chose to believe the conditional 75% of the time, whereas
this percentage was only 37.5% for the low-probability problems. This difference was significant, z = -2.44, p = .015. Similarly, with the MT high-probability problems, subjects chose to believe the conditional 81.3% of the time, whereas for the low-probability problems this occurred only 40.6% of the time, z = -2.997, p = .003. We also used neutral-probability problems close to 50% chance of occurrence to test whether people would in this case guide their belief revision as a function of inconsistencies between mental models. As predicted by the mismatch principle, the conditional was believed significantly more often on the MT problems (71.9%) than on the MP problems (54.7%), z = -1.99, p = .047. However, although in the MT condition the pattern of revision choice was as expected, the conditional being more believed than the categorical, in the MP condition the categorical was not chosen more often (45.3%) than the conditional statement. Figure 4 depicts the mean reaction times for all the conditions.

There was a significant main effect of probability on the reaction times, $F(2, 30) = 128.09, p < .000$. Post-hoc tests indicate that the RT for the neutral-probability problems, collapsed over inference problem, was significantly higher than for the high- and low-probability problems (both p < .000). This might suggest that belief revision is more difficult to perform when the conditional statement has a near 50% chance of occurring. There was also a significant interaction effect between Inference problem and Probability, $F(2, 30) = 3.99, p = .029$. Type of Inference did not affect mean RT times for high- and low-probability conditions, but in the neutral-probability condition subjects had faster decision times for MT problems ($M = 11.08$ s) than for the MP problems ($M = 12.52$ s).

### General Discussion

In both experiments we were able to demonstrate that when the inconsistent set of statements included a high probability conditional, the conditional was believed more than the categorical. Conversely, when the problem set included a low probability conditional, the subjects preferred to believe the categorical statement instead. This was the case for both MP and MT inferences. This shows that people incorporate probability into their belief revision strategy. This highly supports the claims made by advocates of the conditional probability hypothesis in conditional reasoning (e.g. Over, et al., 2007) and shows that this can be extended into the revision process of reasoning problems when inconsistencies arise between statements. These results are related to Politzer and Carles' (2001) finding that the level of probability of conditionals of MP problems affected revision choices (see also Dieusaert, Schaeken, Neys, & d’Ydewalle, 2000).

Also, partial support was found for the mismatch principle (Hasson & Johnson-Laird, 2003; Johnson-Laird, 2006; Johnson-Laird et al., 2004). The conditional was more often believed on the MT neutral-probability problems than on the MP neutral-probability problems. Furthermore, with neutral-probability problems, subjects believed the conditional more often than the categorical statement to a significant degree. This is as the mismatch principle would predict, the mental model of the counterfact $(p)$ matches the first part of the mental model of the conditional $(p q)$. With MP problems, however, the categorical statement was not believed more than the conditional statement, opposing the predicting of the mismatch principle. Keeping in line with this principle, a possible explanation for this could be that people rather prefer to find a mental model within their belief set that matches the mental model of the incontrovertible fact instead of focusing on a mental model that conflicts with that of the given fact. In addition to that, the conditional had a probability close to 50%, the subjects could have also reasoned it was likely possible to occur.

The reaction times in this study also offered some valuable information. In the first experiment, it was shown that the reaction times were faster for high-probability MT problems than for their low probability counterparts. We reasoned that an interaction between the two strategies might have been the underlying cause of this. However, this trend did not emerge in the second experiment. A possible explanation could be that the interaction of the two strategies in the low- and high-probability conditionals...
dampened due to the inclusion of the neutral-probability problems. These might have brought the strategy of finding inconsistencies between mental models more to the surface in this condition. In the second experiment, another interesting finding emerged regarding decision times. Here, the subjects were much faster making their revision choice with high- and low-probability problems than with neutral-probability problems, which portrays the impression that the former two were easier to perform than the latter. Liu et al. (1996) showed that when the conditional had a high perceived sufficiency, subjects correctly concluded that the conclusion, q and not-p on MP and MT problems respectively, follows logically from the premises. With the low perceived sufficiency conditional problem sets, the subjects did not make these correct conclusions. The current reaction time results (in addition to the belief revision responses) further support these and other earlier findings that people reason psychologically. They have a preference for the conditional when it expresses a high probability and a preference for the categorical when the conditional expresses a low probability. This reasoning strategy seems to hamper their performance on reasoning tasks but eases their performance on belief revision tasks.

In conclusion, the two experiments conducted here together showed that people use different reasoning strategies depending on how probable they think the “if p then q” conditional statement is. This shows that people assign probability estimates to the conditionals and illustrates that the probability of the conditional can be regarded as a sort of epistemic entrenchment whereon the belief revision process is based. Moreover, the results of the decision times suggest that it is also easier for people to engage in belief revision when the conditional expresses a clear probability (either high or low) then when it does not and people instead have to rely on mental models to resolve the inconsistency.

Acknowledgments
We thank Susann Ruis and Stefanie Müller for their assistance with the data collection.

References


