The effects of source trustworthiness and inference type on human belief revision

Ann G. Wolf, Susann Rieger, and Markus Knauff

Department of Experimental Psychology and Cognitive Science, Justus Liebig University Giessen, Germany

We investigated whether people revise their beliefs as a function of inference type or source trustworthiness. By doing so we aimed to find out if belief revision is better explained by mental model theory (Johnson-Laird & Byrne, 2002) or by a conditional probability view (Evans, Handley, & Over, 2003; Oaksford & Chater, 2001). We used modified modes ponens (MP) and modus tollens (MT) problems in which the first two premises were uttered by persons with varying degrees of trustworthiness. A third statement was presented as a fact and established inconsistency in the set of propositions. The participants’ task was to indicate which of the first two premises they believed more after receiving the fact. We found that the belief in the conditional premise dropped significantly when this premise was stated by a low- rather than a high-trustworthy source. Moreover we found that the conditional premise was believed more in MT than in MP problems. Both findings are best explained by the conditional probability hypothesis (e.g., Evans et al., 2003).

Keywords: Belief revision, Source trustworthiness, Mismatch principle, Mental models, Conditional probability hypothesis.

Belief revision is the process of changing one’s belief state when a newly acquired fact contradicts one’s beliefs. Belief revision has been a widely studied topic in the areas of artificial intelligence (AI) and philosophy (Alchourrón, Gärdenfors, & Makinson, 1985; Gärdenfors, 1988; Harman, 1986). However, exactly how humans perform belief revision is still a very
unresolved terrain. Human belief revision has typically been investigated with modified modus ponens (MP) and modus tollens (MT) inferences (Evans, Newstead, & Byrne, 1993; Evans & Over, 1996), in which an inconsistency arises between a fact, contradicting the valid conclusion, and the conditional and categorical premises. A reasoner’s task is to decide whether to invest more belief in the first or second premise. Examples of a modified MT problem and a modified MP problem, respectively, are as follows:

If Chris goes to work, then he will take the car
Chris does not take the car
Chris goes to work

If Jane visits a friend, then she brings flowers
Jane visits a friend
Jane does not bring flowers

Research on belief revision has mainly focused on trying to identify a single belief revision strategy (e.g., Byrne & Walsh, 2005, expt 1; Dieussaert, Schaeken, De Neys, & d’Ydewalle, 2000; Elio, 1997, expt 2; Elio & Pelletier, 1997; Politzer & Carles, 2001, expt 1). However, these former studies have not accumulated to a consistent belief revision strategy due to the absence of an overall empirical approach. Our overall goal is to explore a plethora of factors that might affect belief revision. In doing so we aim to test the predictions from mental model theory, a logic-based theory, with those of the conditional probability hypothesis, a probability-based theory. In a recent set of studies (Wolf & Knauff, 2008, 2009) we found that people adjusted their belief in the conditional to the perceived probability of the conditional premise. In the current paper we explore the factors “source trustworthiness” and “inference type”. Social psychology has bundled a great number of studies showing a robust effect of source trustworthiness on persuasion and attitude change (e.g., Eagly, Wood, & Chaiken, 1978; Hovland & Weiss, 1951; Kelman & Hovland, 1953; Mills & Jellison, 1967; Priester & Petty, 1995). Also, source expertise has been shown to influence acceptance rate in deductive reasoning problems (Stevenson & Over, 2001). However, to our knowledge no research has been conducted on the effect of source trustworthiness on reasoning or belief revision. Since variation in the trustworthiness of a statement’s source is expected to affect people’s confidence in the statement, finding such an effect would support probability-based theories. According to mental model theory, people’s reasoning depends on the number of mental models created which in turn is influenced by the type of inference problem. Hence we used variation of inference type to test the predictions of mental model theory.

Mental model theory states that people construct mental models that capture the semantic meaning of premises. Mental models represent true
possibilities of what the premises describe (Johnson-Laird, Girotto, & Legrenzi, 2004). According to mental model theory, people can interpret and represent a conditional in two different ways (Johnson-Laird & Byrne, 2002). The first interpretation conveys a full representation in which only a single explicit model, the \( p \land q \) case, is created and is called the conjunctive response. An additional implicit model indicates other possibilities, although these are ordinarily not fully fleshed out due to working memory limitations. However, when a reasoning task is not too demanding (Johnson-Laird, Byrne, & Girotto, 2009), people are able to assign a material interpretation to a conditional and represent the two further fleshed-out true possibilities in which the antecedent is false, \( \neg p \land q \) and \( \neg p \land \neg q \) (Johnson-Laird & Byrne, 2002). The material interpretation shows that, according to mental model theory, conditionals are always true whenever the antecedent is false. The model \( p \land \neg q \) is considered a false possibility and is therefore not represented.

An important part of the theory concerns the detection of inconsistency. People judge a set of assertions as consistent if they can find a mental model of a possibility in which they are all true. If they fail to do so the set is considered inconsistent. Such judgement usually follows when the conclusion of one’s beliefs does not fit a given fact or observation (Johnson-Laird, Girotto, et al., 2004; Johnson-Laird, Legrenzi, & Girotto 2004; Legrenzi, Girotto, & Johnson-Laird, 2003). Mental model theory proposed the mismatch principle to explain how people reason from inconsistency to consistency (Hasson & Johnson-Laird, 2003; Johnson-Laird, 2006; Johnson-Laird & Byrne, 2002; Johnson-Laird, Girotto, et al., 2004). By this principle the statement that will be discarded or believed less is that statement that has a mental model that mismatches and conflicts with the mental model of the contravening fact. However, this depends on which of the above interpretations is given to the conditional premise. To illustrate with the MP problem, the mental model of the categorical premise is \( p \) and that of the fact \( \neg q \). With a material interpretation of the conditional, the remaining possibilities are fleshed out and people realise that no real inconsistency exists because the model of the fact matches the \( \neg p \land \neg q \) model of the conditional. In that case, the categorical is more likely rejected (Girotto & Johnson-Laird, 2004; Johnson-Laird, Girotto, et al., 2004). However, as previously explained, people tend not to flesh out their mental models into complete models and give a conjunctive response instead. In this regard the conditional premise is represented as a single explicit model. In this case the conditional premise would be abandoned because \( p \land q \) mismatches and conflicts with the model of the fact \( \neg q \); their conjunction would result in the null model. With MT problems the model of the categorical premise is \( \neg q \) and that of the fact \( p \). The more likely conjunctive interpretation of the conditional premise again results in a single \( p \land q \)
model, which is consistent with the fact \( p \). In this case the mismatch principle is taken to imply a preference for retaining the conditional premise over the categorical one. A material interpretation would not affect this preference much (Johnson-Laird, Girotto, et al., 2004).

However, we express some concerns with the mismatch principle in general and with its principles. First, the mismatch principle has undergone very little empirical testing, and mental model theory has received some criticism in recent years in that it does not explain how the content of statements might influence the process of reasoning (e.g., Oberaurer & Wilhelm, 2003). Second, both the conjunctive and the material interpretations of the conditional premise are not unique to the mismatch principle but present an important part of mental model theory in general. These interpretations have been extensively tested, and the results are that there is some modest support for the conjunctive response, but the material interpretation is hardly ever given by participants (Evans et al., 2003; Oberaurer & Wilhelm, 2003; Over, Hadjichristidis, Evans, Handley, & Sloman, 2007). Finally, with a conjunctive interpretation there is a straightforward mismatch between \( \text{not-}q \) and \( q \) in the MP case, but it is hard to understand how \( p \) and \( \text{not-}q \) form a mismatch or conflict in that sense in the MT case.

By mental model theory a conditional is either true or false and conclusions follow deductively from premises that are certain. However, a more recent development in human reasoning states that reasoning is an uncertain process guided by prior belief and knowledge (Evans, 2008; Evans et al., 2003). The main point is that the meaning of a conditional affects how people reason from it. There is a dependency between the antecedent and the consequent in that the truth of the antecedent affects the truth of the consequent (Evans et al., 2003; Oaksford & Chater, 2001, 2007; Oaksford, Chater, & Larkin, 2000). For example, perceived sufficiency and perceived necessity of a conditional's components have shown to influence the acceptance rate of an inference (Bonnefon & Hilton, 2004; Byrne, 1989; Cummins, Lubert, Alksnis, & Rist, 1991; George, 1999; Manktelow & Fairly, 2000; Politzer & Bourmaud, 2002; Stevenson & Over, 1995; Thompson, 1994). This has led to the contemporary view that everyday thought is characterised by probabilistic reasoning whereby people assign degrees of belief, rather than truth values, to conditionals (Evans, 2008; Evans et al., 2003; Evans & Over, 1996; Oaksford & Chater, 2007). Although several probabilistic accounts of reasoning exist, the one we will focus on is the conditional probability hypothesis. The conditional probability hypothesis is inspired by a proposal by the philosopher Ramsey (1990, originally 1929), which is known as the Ramsey test (Evans et al., 2003). Ramsey stated that people judge the degree of probability of the conditional \( \text{if } p \text{ then } q \) as a
degree of belief in \( q \) given \( p \) (Ramsey, 1990, p. 154). The conditional probability hypothesis is a psychological adaptation of his proposal, and explains how the above process is performed mentally and is linked to hypothetical thinking more generally (Evans et al., 2003, 2007; Marcus & Rips, 1979; Over & Evans, 2003; Over et al., 2007). The construction of a conditional, characterised by the word ‘if’, triggers hypothetical thought in which the probability of \( q \) is assessed under the assumption of \( p \) (Evans et al., 2003, 2007; Oaksford & Chater, 2007; Over & Evans, 2003). In this sense the subjective probability of a conditional \( \text{if } p \text{ then } q \) becomes the conditional subjective probability \( P(q|p) \) (e.g., Evans et al., 2007). The judged probability of \( \text{if } p \text{ then } q \) will vary with \( P(q|p) \). To judge \( P(q|p) \), what matters is the probability ratio between the \( p \) \( q \) case and the \( p \rightarrow q \) case (Oberauer & Wilhelm, 2003). The probability of \( p \) \( q \) is compared with the probability of \( p \rightarrow q \). To the extent that \( P(p \ q) \) is higher than \( P(p \rightarrow q) \) then \( P(q|p) \) is high and \( P(\text{if } p \text{ then } q) \) is judged to have a high probability. If, by contrast, \( P(p \ q) \) is judged lower than \( P(p \rightarrow q) \), then \( P(q|p) \) is low and \( P(\text{if } p \text{ then } q) \) is judged to have a low probability (Evans et al. 2003; Over & Evans, 2003). To come back to our MT example at the beginning of the introduction, people can infer upon reading the conditional, e.g., from their experience or a general trend, when leaving for work that the probability of taking the car is higher than the probability of not taking the car. The new information that Chris went to work but did not take his own car could lead people to believe that the car might have broken down. As a broken-down car is an improbable occasion, people will more likely retain their high belief in the conditional premise. Focusing on the hypothetical possibility of \( p \) makes \( \neg p \) cases irrelevant, which is an important distinction from mental model theory (Evans et al., 2003, 2007; Oberauer & Wilhelm, 2003). Over the years researchers have made considerable headway in establishing that people do indeed determine the probability that \( q \) happens on the occurrence of \( p \) (e.g., Evans et al., 2003; Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011; Liu, Lo, & Wu, 1996; Over et al., 2007; Oberauer & Wilhelm, 2003). The relevance of the conditional probability hypothesis has been well recognised and has undergone further development, of which the most notable is the theory of Bayesian rationality by Oaksford and Chater (2007, 2009).

The conditional probability hypothesis was developed to explain uncertain reasoning where one has to develop confidence in the conclusion. In a typical experiment participants rate the truth of an uncertain conditional premise either by assigning probabilities to the four possible conjunctions that match the four rows of a truth table (Oberauer & Wilhelm, 2003; Over et al., 2007), or upon receiving frequency information about the four conjunction cases (Evans et al., 2003). By contrast, in belief revision tasks the
participants receive three statements corresponding to deductive reasoning problems. However, the conclusion is presented as a fact and is the opposite of what is expected to follow from logic or common sense. Even so, we believe that the conditional probability theory is a suitable theory to test reasoning from inconsistency to consistency.

We report three experiments. In each experiment the first two statements were uttered by people whose level of trustworthiness was manipulated. In Experiments 1 and 2 the speakers were represented by their occupations which were rated as high, medium, or low trustworthy. Earlier studies have already shown that people assign different levels of trustworthiness to common occupations (Chun, Campbell, & Yoo, 1975; Rotter & Stein, 1971; Ryckman & Sherman, 1974). In Experiment 3 the premises were uttered by two characters who were described as either apparent trustworthy or untrustworthy. We recorded percentage of belief preference for the conditional premise and the corresponding decision time (DT). We expected that if people reason by the conditional probability hypothesis, they would take into account the level of source trustworthiness. This would translate in a greater preference for the conditional premise over the categorical premise when the former one is uttered by a person whose occupation inspired higher trustworthiness. In turn, the categorical premise would be preferred over the conditional premise if the latter one is uttered by a person whose occupation inspired lower trustworthiness. This reflects the probability ratio between the $p \& q$ case and the $p \& \neg q$ case. On the other hand, if people apply the mismatch principle we expected it to be mirrored by a response asymmetry between inference types. By this route, a clear preference would show up in the MP case and no clear preference for either premise in the MT case, reflecting the conjunctive response. Our third hypothesis concerns the DTs; we expected the DTs to be easier when the level of source trustworthiness was easily discriminable.

EXPERIMENT 1

Method

Participants. A total of 30 participants (20 female) aged 19 to 38 from different faculties from the University of Giessen (also applies to Experiments 2 and 3) took part in the experiment in exchange for a small monetary incentive (4 Euro). None of the participants had pre-knowledge of deductive reasoning and logic (this also applies to Experiments 2 and 3).

Materials and design. A norming study was done to pre-test the trustworthiness of 96 occupations. Booklets were created with general instruc-
tions appearing on the front page and each occupation presented on a separate page. A group of 20 participants rated each job on a 7-point Likert scale from 1 (very untrustworthy) to 7 (very trustworthy). Eight high-judged, eight low-judged, and 16 medium-judged trustworthy occupations were selected for the experiment. The high-trustworthy occupations were rated significantly more trustworthy than the medium-trustworthy and the low-trustworthy occupations, $Z = 3.92$, $p < .0001$, and $Z = 3.93$, $p < .0001$, respectively. The medium-trustworthy occupations were rated significantly more trustworthy than the low-trustworthy ones, $Z = 3.92$, $p < .0001$. From the eight high- and eight low-trustworthy jobs, eight job matches were constructed. In a similar fashion eight job matches were formed from the 16 medium-trustworthy jobs (see Table 1 for an overview of the job matches; in each row the job on the left is matched with the job on the right). MP and MT problems were created in which each job of each match preceded either the conditional or the categorical statement in a counterbalanced manner. The conditional statement always contained the abstract content “If A is the case, then B is the case” in order to avoid possible unwanted influence of content. The abbreviation “high/low” refers to the case in which the conditional was preceded by a high-trustworthy job and the categorical by a low-trustworthy job; the abbreviation “low/high” signals the reversed order.

### Table 1: List of job matches used in Experiment 1 (with reported mean trustworthy ratings)

<table>
<thead>
<tr>
<th>High-trustworthy job</th>
<th>Low-trustworthy job</th>
</tr>
</thead>
<tbody>
<tr>
<td>Police officer ($MN = 5.25$)</td>
<td>Used car salesman ($M = 2.00$)</td>
</tr>
<tr>
<td>Judge ($M = 5.65$)</td>
<td>Spy ($M = 2.60$)</td>
</tr>
<tr>
<td>Pilot ($M = 5.75$)</td>
<td>Salesperson ($M = 3.30$)</td>
</tr>
<tr>
<td>Veterinarian ($M = 5.65$)</td>
<td>Real-estate agent ($M = 2.79$)</td>
</tr>
<tr>
<td>Educator ($M = 5.40$)</td>
<td>Minister ($M = 3.00$)</td>
</tr>
<tr>
<td>Firefighter ($M = 5.78$)</td>
<td>Insurance agent ($M = 2.60$)</td>
</tr>
<tr>
<td>Development aid worker ($M = 5.85$)</td>
<td>Member of the Bundestag (German term for Congress) ($M = 2.70$)</td>
</tr>
<tr>
<td>Psychologist ($M = 5.45$)</td>
<td>Stockbroker ($M = 2.75$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Medium-trustworthy job</th>
<th>Medium-trustworthy job</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bookkeeper ($M = 4.25$)</td>
<td>Banker ($M = 4.43$)</td>
</tr>
<tr>
<td>Photographer ($M = 4.60$)</td>
<td>Hairdresser ($M = 4.40$)</td>
</tr>
<tr>
<td>Businessman ($M = 3.70$)</td>
<td>Accountant ($M = 3.75$)</td>
</tr>
<tr>
<td>Radio presenter ($M = 4.25$)</td>
<td>Travel agent ($M = 4.35$)</td>
</tr>
<tr>
<td>Caretaker ($M = 4.75$)</td>
<td>Student ($M = 4.63$)</td>
</tr>
<tr>
<td>Film producer ($M = 4.20$)</td>
<td>Lawyer ($M = 4.05$)</td>
</tr>
<tr>
<td>Consultant ($M = 3.8$)</td>
<td>Salesperson ($M = 3.75$)</td>
</tr>
<tr>
<td>Hunter ($M = 4.25$)</td>
<td>Butcher ($M = 4.65$)</td>
</tr>
</tbody>
</table>
Medium/medium refers to the case in which both premises are preceded by a medium-trustworthy job. In total, then, 48 inference problems were used for the experiment. A 2 (MP vs MT) × 3 (high-, medium-, low-trustworthiness) within-participant design was used. The dependent variables were the percentage of choosing the conditional premise and the corresponding decision time (DT). The following is an example of a medium/medium MT problem:

Accountant: If A is the case, then B is the case  
Businessman: B is not the case  
Fact: A is the case

Procedure. Participants were seated in a quiet laboratory room in front of a computer. They received verbal and written instructions. They were repeatedly presented with sets of three statements, one at a time. The first two statements were uttered by speakers represented by their occupation and the third statement was presented as a fact. The truth of the first two statements was uncertain but the third statement, causing an inconsistency, was certainly true. The participants’ task was then to resolve this inconsistency by choosing which of the first two statements they believed more. Four practice items preceded the actual experiment. The participants could read the statements in a self-paced manner and press a space-bar-like button on a response box to run through the experiment. After the third statement the participants made their choice by pressing a left or right button. The designation of the left and right button to the conditional and categorical premises was counterbalanced across participants. Furthermore, all 48 problem sets were randomised within participants.

Results

Belief revision choices. The overall mean to believe the conditional was 34.93% ($SD = 14.95$). Belief preferences in the six conditions are depicted in Figure 1 as the mean preference for choosing the conditional premise. An ANOVA revealed significant main effects of Inference problem, $F(1, 29) = 21.70$, $MSE = .096$, $p < .0001$, and Source trustworthiness, $F(1.59, 45.96) = 24.71$, $MSE = .070$, $p < .0001$. The conditional premise was believed more with MT ($M = 45.69$) than with MP problems ($M = 24.17$). With respect to Source trustworthiness, follow-up tests revealed that all three comparisons reached significance: High/low-trustworthy ($M = 50.21$) vs medium/medium-trustworthy ($M = 34.58$) condition, $t(29) = 3.72$, $p = .001$; High/low-trustworthy vs low/high-trustworthy ($M = 20.00$) condition, $t(29) = 5.81$, $p < .0001$; Medium/medium-trustworthy vs low/high-trustworthy condition, $t(29) = 4.46$, $p < .0001$. The Inference problem × Source trustworthiness interaction did not reach significance, $F(2, 58) < 1$.  

[424 WOLF, RIEGER, KNAUFF]
Figure 1. Mean percentages of belief preference for the conditional premise in each condition in Experiment 1. MP_HTrust, MP_MTrust, MP_LTrust = modus ponens high-trustworthy, medium-trustworthy, and low-trustworthy inference problems, respectively. MT_HTrust, MT_MTrust, MT_LTrust = modus tollens high-trustworthy, medium-trustworthy, and low-trustworthy inference problems, respectively.

Figure 2. Mean decision times in seconds for the conditional premise in each condition in Experiment 1.

Decision times. The overall DT was 6.54 s. (SD = 1.80). Figure 2 depicts the mean DTs for the six conditions. A repeated-measures ANOVA revealed significant differences due to Inference problem, $F(1, 29) = 27.33$,
MSE = 2.02, p < .0001 and Source trustworthiness, F(1.14, 32.95) = 214.96, MSE = 9.50, p < .0001. The DTs were significantly higher for MT (M = 7.09) than for MP problems (M = 5.99). As for Source trustworthiness, all three comparisons were significant: High/low-trustworthy (M = 3.52) vs medium/medium-trustworthy condition (M = 11.59), t(29) = 16.25, p < .0001; High/low-trustworthy vs low/high-trustworthy condition (M = 4.52), t(29) = 6.50, p < .0001; Medium/medium-trustworthy vs low/high-trustworthy condition, t(29) = 13.61, p < .0001. The interaction between Inference problem and Source trustworthiness was also significant, F(1.46, 42.25) = 12.73, MSE = 2.39, p < .0001. With MT problems, the DTs were significantly higher when the conditional premise was coupled with a low-trustworthy occupation (M = 3.47), t(29) = 8.88, p < .0001. Also the DTs in both high/low- and low/high-trustworthy conditions were significantly lower than in the medium/medium-trustworthy condition (M = 12.14), t(29) = 14.83, p < .0001, and t(29) = 11.21, p < .0001. With MP problems, the DTs in the high/low condition (M = 3.57) and the low/high condition (M = 3.36) were significantly lower than in the medium/medium-trustworthy condition (M = 11.04), t(29) = 12.53, p < .0001, respectively.

Discussion

Source trustworthiness served as a predictor of belief revision evidenced by higher confidence in the conditional premise when this premise was uttered by a high-trustworthy speaker, confirming our first prediction. In line with our second prediction, inference type also influenced belief revision. With MP problems, the conditional gained little belief and with MT problem the belief in the conditional approximated chance level. This finding mirrors the conjunctive interpretation and is supportive of the mismatch principle. Finally, the DTs showed that a clear contrast in levels of source trustworthiness eased performance, confirming our third hypothesis. Thus both the belief revision responses and DTs showed the effects of source trustworthiness and inference type on belief revision.

EXPERIMENT 2

Experiment 1 showed a clear effect of both source trustworthiness and inference type on belief revision. However, even though our aim in Experiment 1 was to investigate whether there is any effect of trustworthiness, using content-rich statements would more closely resemble real-life situations. In Experiment 2 we included statements that are likely to be uttered in a conversation between two persons that represent the two occupations of each match.
Method

Participants. A total of 36 participants (24 female) aged 19 to 24 ($M = 23.19, SD = 3.42$) took part in the experiment in exchange for a small monetary incentive (4 Euro).

Materials and design. Four high-low/low-high, and four medium-medium trustworthy job matches were taken from the pool of job matches used in Experiment 1. For each match six conditional statements were constructed. In addition each of the six conditionals was preceded by each of the two jobs of each match. The resulting 96 conditionals were used for a norming study in which booklets were used with each statement appearing on a separate page, in random order. A total of 20 participants rated under each statement, on a rating scale ranging from 0% (very unlikely) to 100% (very likely), how likely they thought it that the statement was true. From each job match the conditional that was rated around 50% with either job was taken for the experiment. That is, the one conditional from each of the eight job matches was taken twice, each one preceded by one of the two jobs. A Friedman Test showed that all 16 conditionals did not significantly differ from one another with respect to probability, $\chi = 16.67, p = .339$. Wilcoxon Signed Rank tests were performed to check if the conditionals per match had equal or similar probability. No significant differences showed up (all $p$-values were above .05). For the experiment, MP and MT problems were created whereby each job of each match preceded the conditional premise and categorical premise in a counterbalanced fashion (medium1-medium2 and medium2-medium1 serve as abbreviations for the counterbalancing of the medium-trustworthy job matches). This resulted in 32 inference problems. The design was identical to that of Experiment 1. Appendix A lists the conditionals used with the accompanying job matches. The following is an example of a high/low trustworthy MP problem:

Police officer: If Karl sells a damaged car, then he offers the car for a reduced price.
Used car salesman: Karl sells a damaged car.
Fact: Karl does not offer the car for a reduced price

Procedure. The procedure was identical to that of Experiment 1.

Results

Belief revision choices. Paired-samples $t$-tests performed on both MP and MT medium-trustworthy problems showed that it made no difference which occupation of each match preceded the conditional or the categorical premise, $t(35) < 1$ and $t(33) < 1$, respectively. For this reason we collapsed
over medium1/medium2 and medium2/medium1 inference problems and combined them, for both MP and MT problems. The overall preference for the conditional premise was 42.27% ($SD = 17.32$). Mean percentages of belief revision choices in the six conditions are depicted in Figure 3. A repeated-measures ANOVA revealed a main effect of Inference problem, $F(1, 35) = 101.69$, $MSE = .087$, $p < .0001$, such that the conditional premise was endorsed significantly more with MT ($M = 62.62$) than with MP problems ($M = 22.11$). No main effect of Source trustworthiness emerged, $F(2, 70) < 1$. Also, no Inference problem $\times$ Source trustworthiness interaction was detected, $F(2, 70) = 1.21$, $p = .304$.

*Decision times.* Also with DTs, for both MP and MT problems it made no significant difference which of the two medium-trustworthy occupations preceded which statement, $t(35) = 1.99$, $p = .056$ and $t(35) = .342$, $p = .73$, respectively. Therefore the two conditions of each inference problem type were collapsed and from here on are referred to as MP medium-trustworthy and MT medium-trustworthy problems. The overall DT was 5.74 s ($SD = 1.62$). Figure 4 shows the mean decision times for the six conditions. A repeated-measures ANOVA showed no main effects of Inference problem, $F(1, 35) = 1.46$, $MSE = 2.08$, $p = .236$, and Source trustworthiness, $F(2, 70) = 1.30$, $MSE = 2.44$, $p = .280$. However, an Inference problem $\times$ Source trustworthiness interaction emerged, $F(1.589, 55.627) = 6.12$,
\(MSE = 1.82, p = .007\). For MT problems the DTs were significantly lower when the conditional premise was coupled with a high-trustworthy \((M = 5.08)\) or a low-trustworthy occupation \((M = 5.37)\) than with a medium one \((M = 6.23)\), \(t(35) = 3.37, p = .002\), and \(t(35) = 2.24, p = .032\), respectively. In contrast to Experiment 1, coupling the conditional premise with a high-trustworthy occupation rather than a low one did not significantly influence DTs, \(t(35) > 1\). For MP problems, the DTs for choosing the conditional premise were similar regardless the speaker’s trustworthiness.

**Discussion**

The effect of source trustworthiness on belief revision disappeared when the premises contained meaningful content that could relate to a possible conversation. Instead, belief revision was solely guided by inference type. However, this confirmed our second prediction only partially. Again, an asymmetry response indicated a conjunctive interpretation. The conditional premise gained little belief with MP problems; however, with MT problems there was a clear preference for the conditional premise. Finally, with MT problems, the participants were quicker to respond when there was a clear contrast between trustworthiness of the two speakers than when the speakers were equally trustworthy. This partially confirms our third prediction.

The lack of impact of source trustworthiness is rather surprising. Although we performed the norming study for the current experiment to
establish equal plausibility, it was expected that when two occupations with contrasting levels of trustworthiness appeared together, people would make a relative judgement and prefer the statement that is preceded by the more-trustworthy occupation. However, the findings suggest that the effect of source trustworthiness functions through the believability it gives to the statements.

Therefore in the third experiment we also used medium-probability statements but the probability of these statements was pretested without a source. A personality description of two different persons, given prior to the experiment, reflected the sources’ trustworthiness.

**EXPERIMENT 3**

**Method**

**Participants.** A total of 30 participants (21 females) aged 19 to 30 ($M = 23.10, SD = 3.08$) took part in the experiment in exchange for a small monetary incentive (6 Euro).

**Materials and design.** The experiment utilised a 2 (Inference problem: MP vs MT) × 2 (Trustworthiness of source: trustworthy vs untrustworthy) within-participants design. The dependent variables were again belief preference for the conditional premise and corresponding DTs. Prior to the start of the experiment, participants were presented with a trustworthy (Christian) and an untrustworthy (Mark) person description:

Christian is known as a sincere, honest, reliable, responsible, and caring person. Christian is 45 years old and has had the same job for 20 years. He has been married now for 15 years. Christian is described as having a big heart and someone whom you can always rely on.

Mark is known as a reserved, dishonest, unreliable, and uncaring person without any feeling of responsibility. Mark is 45 years old and has been fired from several jobs due to fraudulent activities. Mark’s wife divorced him after five years of marriage, because he had repeatedly cheated on her. Mark is described as selfish and someone you cannot rely on.

Six conditionals that received a 50% probability rating in previous experiments were used in the current experiment (Wolf & Knauff, 2008). These six conditionals, presenting daily actions, did not differ from each other with respect to probability, $\chi = 5.40, p = .369$. A total of 24 inference problems were presented consisting of the following four conditions: Six high/low and six low/high MP problems, and six high/low and six low/high
MT problems. High/low indicates that the conditional was uttered by Christian and the categorical by Mark, and low/high means the reverse. The experiment consisted of six blocks, each containing four inference problems, one problem of each condition. Appendix B lists the conditional statements used for the experiment. An example of a low/high MP problem is as follows:

Mark: If Katharina cooks in the evening, then she prepares a rice dish  
Christian: Katharina cooks in the evening  
Fact: Katharina does not prepare a rice dish

The order of the problems was randomised per participant. To keep the participants aware of the statements’ sources, additional information was offered concerning the sources’ trustworthiness after each four problems (see Appendix B). The order of these six pairs of statements was fixed across participants.

Procedure. The procedure was identical to that of Experiments 1 and 2 except that at the end of the experiment the participants were asked to rate the trustworthiness of both sources on a 7-point Likert scale.

Results

Belief revision choices. The overall preference for the conditional premise was 41.25% \( (SD = 14.16) \). The mean belief preferences per condition are shown in Figure 5. An ANOVA showed a main effect of Inference problem, \( F(1, 29) = 34.70, MSE = .079, p < .0001 \) such that the conditional statement was endorsed more for MT \( (M = 56.39) \), than for MP problems \( (M = 26.11) \). A main effect also showed up for Source trustworthiness, \( F(1, 29) = 29.03, MSE = .055, p < .0001 \), in that the conditional premise was believed more when it was uttered by the high-trustworthy source \( (M = 52.78) \) than by the low-trustworthy source \( (M = 29.72) \). The Inference problem \( \times \) Source trustworthiness interaction also led to significance, \( F(1, 29) = 5.48, MSE = .022, p = .026 \). For both MP and MT problems, the belief in the conditional premise was greater when it was preceded by a high-trustworthy than a low-trustworthy source. However, this difference was much greater for MT \( (M = 71.11 \text{ vs } M = 41.67) \); \( t(29) = 5.23, p < .0001 \), than for MP problems \( (M = 34.44 \text{ vs } M = 17.78) \); \( t(29) = 3.75, p = .001 \).

A paired-samples \( t \)-test showed that the high-trustworthy source (i.e., Christian) had a significantly higher mean trustworthy rating \( (M = 5.65, SD = 1.09) \) than the low-trustworthy source (i.e., Mark; \( M = 2.60, SD = 1.07) \); \( t(29) = 8.31, p < .0001 \). This shows that the overall higher belief in the
conditional premise when it was preceded by a high- rather than low-trustworthy source is indeed attributable to the effect of source trustworthiness.

**Decision times.** The overall DT was 13.23 s ($SD = 4.78$). Figure 6 depicts the mean DTs for the four conditions. A $2 \times 2$ analysis of variance yielded one significant effect, that of Source trustworthiness, $F(1, 29) = 62.50$, $MSE = 15.79$, $p < .0001$. Deciding which statement to believe more took longer when the conditional premise was uttered by the low-trustworthy ($M = 16.10$) than by the high-trustworthy source ($M = 10.36$).

**Discussion**

The results of Experiment 3 are similar to those of Experiment 1. That is, source trustworthiness had an influence on belief revision such that it profoundly lowered the belief in the conditional premise when this premise was coupled with a person of low rather than high trustworthiness. The DTs also mirrored the influence of source trustworthiness in that belief preference for the conditional premise seemed an easier choice when it was coupled with a high- rather than a low-trustworthy source. Again, the mismatch principle received only partial support. The conditional premise gained little confidence with MP problems, reflecting the mismatch principle. However, with MT problems the conditional premise was believed
more in the high-trustworthy condition than in the low-trustworthy condition.

**GENERAL DISCUSSION**

This study presented participants with MP and MT problems with varying degrees of source trustworthiness on which they had to perform belief revision. The goal was to contrast two different accounts of belief revision: mental model theory and the conditional probability hypothesis. Source trustworthiness was expected to function as an uncertainty factor by which people judge their confidence in a statement. Following the main idea of the conditional probability hypothesis, higher belief would then be invested in the premise uttered by the more trustworthy person. On the other hand, resolving belief revision by means of the mismatch principle would be apparent by an inference type asymmetry.

The results from Experiments 1 and 3 clearly show that source trustworthiness influenced belief revision, confirming our first hypothesis. Overall, the conditional premise was believed more when it was uttered by a high-trustworthy speaker than by a low-trustworthy speaker. The preference for the conditional premise over the categorical one in this context shows that the subjective probability of \( p \& q \) was judged higher than the subjective probability of \( p \& \text{not-}q \). This is also in synchrony with theories stating that prior world knowledge is integrated in reasoning performance (Evans, 2008; Evans et al., 1993). We also found an MP–MT asymmetry in belief revision. In all three experiments preference for the conditional
premise was found to have its lowest percentage with the MP relative to the MT problems. This finding supports the main idea of the mismatch principle (Johnson-Laird & Byrne, 2002), as there was a clear mismatch between the model of the fact and the conditional’s explicit model in MP problems. Nonetheless, the mismatch principle falls short of explaining the results fully on two accounts. First, with MT problems it was expected that no clear preference for either premise would show up due to a lack of a clear mismatch. However, the conditional premise was believed more than the categorical premise. This seemed to be most likely caused by a clear match between the model of the fact and the conditional’s mental model. Perhaps the solution in the case where there is no obvious mismatch is to find the best-fitting match. In any case, mental model theory should clarify what is meant by a mismatch and how the mismatch principle applies to MT problems.

Second, mental model theory dictates that the explicit mental model is elicited much more easily than a conditional’s other possible models. This seems indeed to have been the case in the current experiments. However, mental model theory has extended its theory with “pragmatic modulation” to deal with content effects (Johnson-Laird & Byrne, 2002; Johnson-Laird, Legrenzi et al., 2004). Pragmatic modulation implies that general knowledge and the circumstances in which a conditional premise is uttered can modulate the interpretation of a conditional premise. As a consequence pragmatic modulation can hamper, but also motivate, the construction of further models. This implies that a conditional premise that carries low confidence should trigger the process of pragmatic modulation, which in turn should elicit the remaining models not-p & q and not-p & not-q. In the MP case it would then become apparent that the model of the fact is represented in the not-p & not-q model of the conditional premise. Instead we actually found the opposite, especially in Experiment 3. The interaction effect showed that in both inference problems the belief in the conditional was lower when it was expressed by a low-trustworthy person. Mental model theory should extend its mismatch principle to integrate these points.

We will consider now how the conditional probability hypothesis would explain the observed asymmetry. When judging a conditional probability, people make a subjective probability estimate that q is the case upon assuming p. New information presented as a fact can make the probability of the conditional premise go up or down (Evans et al., 2003; Fugard et al., 2011; Oaksford et al, 2000; Oberaurer & Wilhelm, 2003; Over & Evans, 2003; Over et al., 2007). In the MP case, asserting with certainty that the conclusion \( P(\text{not-q}) = 1 \) implies that \( P(q) = 0 \) and so \( P(q|p) = P(\text{if } p \text{ then } q) = 0 \). In the MT case, asserting p with certainty makes \( P(p) = 1 \) which implies that \( P(q|p) = P(\text{if } p \text{ then } q) = q \). This leads to two consequences. First, belief in the conditional premise for MP is very low and, if not zero, it
will in any case always be lower than it is for MT (unless that one learns that not-\(q\) is certain). In this case the conditional probability hypothesis can explain the asymmetry more adequately and is as such a direct challenge to the mismatch principle as explained by mental model theory. Second, the effect of manipulating trustworthiness of the categorical premise should be greater for MT than for MP problems. To explain, the categorical premise will exert little influence whenever the consequent is null and the conditional premise approaches zero, which is the case with MP problems. On the other hand, the probability of the conditional premise will approximate that of the consequent whenever the belief in the antecedent is 1. In other words, the belief in the relationship between \(p\) and \(q\) becomes primarily a matter of how much belief one has in the consequent given the antecedent, since the antecedent is certain. For example, if the categorical not-\(q\) is uttered by a low-trustworthy person, this will lower the belief in the conditional (Oaksford & Chater, 2007, 2009; Oaksford et al., 2000).

In short, the current results indicate that the conditional probability hypothesis is better in explaining both context effects, in this case speakers’ level of trustworthiness, and the MP–MT asymmetry than mental model theory. We do not mean to rule out the role of the mismatch principle in belief revision. However, mental model theory should extend its mismatch principle to cases in which a set of statements is not equally probable. Moreover, although we found no support for the fleshing out of further models, mental model theory should clarify how the material implication and pragmatic modulation relate to each other and how they operate in belief revision. Even though we found support for the conditional probability hypothesis, it can be tested more thoroughly in belief revision. For example, participants could explicitly indicate their initial conditional probability and this prior estimate can be compared to their adjusted estimate after receiving the new information. Such paradigm will also more closely resemble prior work on the conditional probability hypothesis (Evans et al., 2003; Fugard et al., 2011; Oberaurer & Wilhelm, 2003; Over et al., 2007).

Our second expectation was that belief revision would be fastest when the two uncertain premises were uttered by people of different source trustworthiness. What we found was that this was only the case when such a condition was contrasted with instances in which the statements were preceded by sources of equal trustworthiness. This illustrates that the ease of relying on source information in belief revision is much dependent on the social context. This finding relates to studies showing that perceived knowledge about or attitudes towards a person is influenced by information or attitude of other people encountered in close temporal proximity (Markman & McMullen, 2003; Mussweiler, 2003; Tormala & Petty, 2007). The applied standard of comparison in Experiment 1 was that of a “clear difference” (High/low and Low/high) versus “no clear difference” (medium/medium).
In contrast the standard of comparison in Experiment 3 was “High” versus “Low”. This finding is important because people rarely ever receive information in isolation, but rather from multiple sources within a single context.

In short, we found that the propensity to believe one statement over the other varied as a function of source trustworthiness as well as inference type. Furthermore, the findings on source trustworthiness show the importance of social context effects in belief revision. Additional knowledge and belief define for a large part how conditionals are understood in a certain context (Oaksford & Chater, 1991; Over & Evans, 2003). The current results show that defeasibility in pragmatics extends to belief revision and this seems to be best explained by the conditional probability hypothesis. Finally, future research would profit from finding out what other factors from the different disciplines of psychology affect belief revision.

REFERENCES


APPENDIX A

List of conditionals with their trustworthy job matches used in Experiment 2

<table>
<thead>
<tr>
<th>Conditional</th>
<th>Job matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>If Karl sells a damaged car, then he offers the car for a low prize</td>
<td>Policeman vs used car salesman</td>
</tr>
<tr>
<td>If Sophia enters a company, then she wears an access badge</td>
<td>Judge vs spy</td>
</tr>
<tr>
<td>If Anne books a flight, then she chooses the flight company with the best reputation</td>
<td>Pilot vs salesperson</td>
</tr>
<tr>
<td>If Lukas owns a fire extinguisher, then he also has a smoke detector</td>
<td>Fire-fighter vs Insurance agent</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conditional</th>
<th>Job matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>If Laura buys an item, then she pays with credit card</td>
<td>Bookkeeper vs Banker</td>
</tr>
<tr>
<td>If Paul seeks financial help, then he consults the yellow pages</td>
<td>Businessman vs accountant</td>
</tr>
<tr>
<td>If Nils hires an actor for a movie, then he pays the actor his share of the profit</td>
<td>Film producer vs lawyer</td>
</tr>
<tr>
<td>If Nicole owns a farm, then she gives her animals special diets</td>
<td>Hunter vs butcher</td>
</tr>
</tbody>
</table>

APPENDIX B

List of conditionals used in Experiment 3

<table>
<thead>
<tr>
<th>Conditional</th>
</tr>
</thead>
<tbody>
<tr>
<td>If Katharina cooks in the evening, then she prepares a rice dish</td>
</tr>
<tr>
<td>If Hugo is allergic to fur-bearing animals, then he has/gets fish</td>
</tr>
<tr>
<td>If Bruno is hungry, then he just eats a cracker</td>
</tr>
<tr>
<td>If Hendrik is sick, then he will go to work</td>
</tr>
<tr>
<td>If Peter leaves his house, then he takes the backdoor</td>
</tr>
<tr>
<td>If Karl goes to work, then he takes the car</td>
</tr>
</tbody>
</table>
List of additional information concerning the sources used in Experiment 3

Each of the following additional pairs of statements was preceded by the commentary: “Now you receive the additional information”

Christian finds it important to respect norms and values
Mark plays people against each other for his own gain
Christian has received a medal for community service
Mark received a court summons due to violation of privacy
Christian is always eager to donate money to charity
Mark gambles money that he receives from friends
Christian tells his honest opinion without insulting anyone
Mark likes to lie to people for personal gain
Christian does his assignments at work very conscientiously, while
Mark often does his work assignments negligible
Christian commits to appointments and promises, while
Mark often cancels appointments and makes false promises