Interpretation of and Reasoning with Conditionals - Probabilities, Mental Models, and Causality

Dissertation Andrea Weidenfeld

vorgelegt zur Erlangung des akademischen Grades der Doktorin der Philosophie (Dr. phil.)

eingereicht bei der Humanwissenschaftlichen Fakultät der Universität Potsdam im Fach Psychologie im Jahr 2003

Abstract

In everyday conversation "if" is one of the most frequently used conjunctions. This dissertation investigates what meaning an everyday conditional transmits and what inferences it licenses. It is suggested that the nature of the relation between the two propositions in a conditional might play a major role for both questions. Thus, in the experiments reported here conditional statements that describe a causal relationship (e.g., "If you touch that wire, you will receive an electric shock") were compared to arbitrary conditional statements in which there is no meaningful relation between the antecedent and the consequent proposition (e.g., "If Napoleon is dead, then Bristol is in England").

Initially, central assumptions from several approaches to the meaning and the reasoning from causal conditionals will be integrated into a common model. In the model the availability of exceptional situations that have the power to generate exceptions to the rule described in the conditional (e.g., the electricity is turned off), reduces the subjective conditional probability of the consequent, given the antecedent (e.g., the probability of receiving an electric shock when touching the wire). This conditional probability determines people's degree of belief in the conditional, which in turn affects their willingness to accept valid inferences (e.g., "Peter touches the wire, therefore he receives an electric shock") in a reasoning task. Additionally to this indirect pathway, the model contains a direct pathway: Cognitive availability of exceptional situations directly reduces the readiness to accept valid conclusions.

The first experimental series tested the integrated model for conditional statements embedded in pseudo-natural cover stories that either established a causal relation between the antecedent and the consequent event (causal conditionals) or did not connect the propositions in a meaningful way (arbitrary conditionals). The model was supported for the causal, but not for the arbitrary conditional statements. Furthermore, participants assigned lower degrees of belief to arbitrary than to causal conditionals. Is this effect due to the presence versus absence of a semantic link between antecedent and consequent in the conditionals?

This question was one of the starting points for the second experimental series. Here, the credibility of the conditionals was manipulated by adding explicit frequency information about possible combinations of presence or absence of antecedent and consequent events to the problems (i.e., frequencies of cases of 1. true antecedent with true consequent, 2. true antecedent with false consequent, 3. false antecedent with true consequent, 4. false antecedent with false consequent). This paradigm allows furthermore testing different approaches to the meaning of conditionals (Experiment 4) as well as theories of conditional reasoning against each other (Experiment 5).

The results of Experiment 4 supported mainly the conditional probability approach to the meaning of conditionals (Edgington, 1995) according to which the degree of belief a listener has in a conditional statement equals the conditional probability that the consequent is true given the antecedent (e.g., the probability of receiving an electric shock when touching the wire). Participants again assigned lower degrees of belief to the arbitrary than the causal conditionals, although the conditional probability of the consequent given the antecedent was held constant within every condition of explicit frequency information. This supports the hypothesis that the mere presence of a causal link enhances the believability of a conditional statement. In Experiment 5 participants solved conditional reasoning tasks from problems that contained explicit frequency information about possible relevant cases. The data favored the probabilistic approach to conditional reasoning advanced by Oaksford, Chater, and Larkin (2000).

The two experimental series reported in this dissertation provide strong support for recent probabilistic theories: for the conditional probability approach to the meaning of conditionals by Edgington (1995) and the probabilistic approach to conditional reasoning by Oaksford et al. (2000). In the domain of conditional reasoning, there was additionally support for the modified mental model approaches by Markovits and Barrouillet (2002) and Schroyens and Schaeken (2003). Probabilistic and mental model approaches could be reconciled within a dual-process-model as suggested by Verschueren, Schaeken, and d'Ydewalle (2003).

Contents

1 Introduction

"If you touch that wire, you will get an electric shock". This conditional statement might be uttered by Paul as a warning to his friend Steve. Taking the statement for true, Steve does not touch the wire and does not get a shock. One question addressed in this dissertation is whether under these circumstances, Paul's conditional statement will be considered as true or false. Confronted with this question, Steve reckons that he will never know for sure whether the statement is true unless he (or somebody else) touches the wire and either receives an electric shock or does not. In the first case, the conditional is true; in the second case it is false. Certainly, Steve prefers to circumvent this option, since he does not want to take the risk of receiving an electric shock. Instead, Steve might come to the conclusion that whether the statement is true depends on whether the wire is dead or alive. Because if the wire was dead, touching it would not lead to an electric shock. Steve does not know whether the wire is dead, therefore he reaches the conclusion that Paul's statement is *probably true*. In everyday life sentences are often not just true or false, but instead can be trusted with a certain degree of confidence. This intuition that certainty comes in degrees has been captured by a philosophical theory about the meaning of conditionals that Edgington (2003) calls suppositional theory. According to the suppositional theory the addressee of a conditional statement supposes that the antecedent p is true and considers what he thinks about the likelihood of the consequent q under that supposition. Edgington (2003) argues that a main purpose for the existence of indicative conditionals expressing beliefs is the possibility to express uncertainty and to think about possible consequences of likely as well as unlikely events, e.g. "If it rains tomorrow, I won't visit my friend" or "If I win the lottery, I will buy a sailing boat". Furthermore, statements are often helpful even if they don't express certainty, for instance when I ask my doctor "Will I survive the operation? Will the operation cure me?" or my estate agent "Will I be able to sell my house within three months for 300,000 euros?". Usually, suitable answers will be expressing degrees of confidence, but no definite Yes-or-No-answer, e.g., "Odds are 90 to 10 that you will survive the operation, and 95 to 5 that the operation will cure you" or "It's likely that you will be able to sell your house for that price, but it might take a few more months than three". Suppositional theory takes uncertainty seriously. According to the approach a conditional will be considered as true respectively highly believable if the probability of the consequent q given the antecedent p, P(consequent | antecedent), is high. This formalization of the suppositional theory has been called the *conditional probability approach* to the meaning of conditionals and has been recently been supported empirically in psychological experiments (Evans, Handley, & Over, 2003; Oberauer & Wilhelm, 2003a).[1] The present

[1] Though there were no systematic empirical investigations on that problem yet, psychologist had previously already *suspected* that conditionals in everyday conversations "normally express uncertain or probabilistic relationships rather than factual ones" (Newstead, Ellis, Evans, & Dennis, 1997, p. 72). Newstead et al. (1997, p. 72) continued to suppose that "this uncertainty comes in two forms: first there is uncertainty about the extent to which the antecedent implies that the consequent will occur; and second there is uncertainty as to the extent to which the occurrence of the consequent implies that the antecedent must have occurred previously". The first uncertainty corresponds to the conditional probability of the consequent given the antecedent, $P(q|p)$, while the latter matches the reverse probability of the antecedent given the consequent, $P(p|q)$.

dissertation will add further support to this theory, while nevertheless challenging the view that the meaning of conditionals can be completely reduced to conditional probabilities. Consider for example these two conditionals:

1. If the church bells are ringing in Manchester,
then the workmen in *Manchester* knock off work.

2. If the church bells are ringing in Manchester,
then the workmen in *London* knock off work.

Even if both sentences shared the same conditional probability of the consequent given the antecedent (which would happen to be the case if the workmen in Manchester and in London knocked off work at the same time), there is reason to suppose that the believability of the statements is nonetheless different. Because there is a meaningful relation between antecedent and consequent in the first sentence that is absent in the second one – there is no obvious reason that connects ringing bells in Manchester with knocking off work in London. Edgington (1995, pp. 259-269) discusses that many philosophers consider conditionals in which antecedent and consequent are irrelevant for each other like "If Napoleon is dead, then Bristol is in England" as not acceptable or even false. She herself argues that a conditional is *misleading* if the consequent q is as likely under p as it is under *not-p* ($\neg p$). In this case $P(q|p)$ equals $P(q|\neg p)$. Imagine, for example that I know that a football match will be cancelled because several players have the flu. The conditional "If it rains, the match will be cancelled" and the conditional "If it doesn't rain, the match will be cancelled" share a conditional probability. Uttering either conditional will be misleading in this situation, because the match will be cancelled anyway and not because of or depending on the rain. Although Edgington (1995) brings up these considerations on the potential relevance of the presence versus absence of a semantic relation between antecedent and consequent, she does not suggest how to integrate this into the suppositional theory. From a psychological viewpoint it is a very interesting question whether the presence of a semantic link plays a role for the meaning of a conditional. Is it a necessary condition for each conditional - as suggested by philosophers? Alternatively, it might not be a necessary condition, but it may increase the degree of belief with which people trust a conditional – this *semantic link hypothesis* will be investigated empirically in this dissertation.

In everyday conversation conditional statements often imply that certain inferences are licensed while others are not. If for example a mother advises her daughter in the morning with the conditional statement "If you are late for supper, you won't get desert", the daugther will take for granted that she gets a desert if she arrives on time for supper – an inference that is logically not necessarily valid, but which has most likely indeed been implied by her mother. Geis and Zwicky (1971) have called such conclusions "invited inferences" (see also Fillenbaum, 1977; Grice, 1975; Sperber & Wilson, 1995). By the same token, in everyday conditionals often even valid inferences are not implied by the speaker. From the warning "If you don't dry you hair after swimming, you will get a flu" and the knowledge that the addressee did not get a flu, it seems unwarranted to conclude that he actually followed the advice and dried his hair after swimming. The meaning of conditional statements is intimately connected with the question what kind of inferences are licensed by the statements, this dissertation will therefore investigate the meaning and reasoning from conditionals simultaneously.

Semantic links come in many flavors: There are for example promises and warnings (e.g. "If you clean up your room, we will go to the zoo on Sunday" and "If you don't clean up your room, we won't go to the zoo on Sunday"), tips ("If you clean up your room, Dad will go to the zoo with you on Sunday"), deontic links ("If you want to enter Australia as a tourist, then you must have a valid visa"), links concerning social exchange ("If you clean up the dishes, I will mop the floor") and causal links ("If you cut your finger, then you will bleed"). This list is only an illustration and by no means exhaustive.

For this dissertation, causal relations were picked out of all the possible candidates to start exploring the role of semantic links in conditionals. Statements and reasoning about causal relations are very common in daily life. Understanding causal relations is crucial since it allows predicting and controlling effects of actions, thus helping to achieve personal goals. Reasoning with natural causal conditional statements like "If you cut your finger, then you will bleed" or "If you fertilize a flower, then it will bloom" has been the matter of intense theoretical and empirical work in the last decade. In the next section, these theoretical and empirical achievements will be integrated with the conditional probability approach to the meaning of conditionals to develop an integrated model of understanding and reasoning from causal conditionals. In the first experimental series (Experiments 1 to 3), this integrated model will be tested for two sorts of causal conditionals as well as for noncausal conditionals. The second experimental series (Experiments 4 and 5) investigates whether the presence of an explicit causal link plays a role for the believability of a conditional and the reasoning from these conditionals even if the conditional probability, $P(q|p)$, is controlled experimentally.

2 A Model for Causal Conditionals

Natural causal conditionals like "If you fertilize a flower, then it will bloom" differ in the number and availability[2] of *disabling conditions* that might prevent the effect to occur in the presence of the cause, and of *alternative causes* that can bring about the effect in the absence of the cause. For "If you fertilize a flower, then it will bloom" it is quite easy to imagine several disabling conditions (e.g., the flower has not been watered enough, the flower suffers from varmints, the flower grows in infertile ground). Whereas there is almost no disabling condition for "If you cut your finger, then you will bleed", because nearly nothing prevents a finger from bleeding after it has been cut. In a variety of experimental investigations using natural causal conditionals with forward causality (i.e., if cause, then effect), it has been shown that that the availability of disabling conditions reduces people's readiness to accept the logically valid inference forms modus ponens (MP: "if *p* then *q*, *p*, therefore *q*") and modus tollens (MT: "if *p* then *q*, *not-q*, therefore *not-p*") (Cummins, Lubart, Alksnis, & Rist, 1991; Cummins, 1995; Quinn & Markovits, 1998; Thompson, 1994): Participants accepted fewer MP and MT inferences from conditionals with many

---

[2] *Availability* refers to the ease with which many different situations can be brought to mind.

disabling conditions than from conditionals with few disabling conditions: Availability of disabling conditions suppresses MP and MT in causal conditionals. The effect of availability of disabling conditions for MT even persists with instructions that put heavy emphasis on the logical nature of the task (Vadeboncoeur & Markovits, 1999). The inferences acceptance of the consequent (AC: "if *p* then *q*, *q*, therefore *p*") and denial of the antecedent (DA: "if *p* then *q*, *not-p*, therefore *not-q*"), which are not valid according to standard propositional logic, have been shown to be suppressed by the availability of alternative causes. The integrated model of interpretation of and reasoning from causal conditionals that will be introduced in the following captures simultaneously causal forward conditionals (if effect, then cause) as well as causal backward conditionals (if cause, then effect). To achieve this goal, the concept of *exceptional situations* has to be introduced: Exceptional situations are circumstances that have the power to generate exceptions to the conditional rule, that is, cases in which the antecedent event is true but the consequent is false. For conditionals with forward causality, exceptional situations constitute disabling conditions. For conditionals with a backward causal link, exceptional situations correspond to alternative causes.

In the model, two causal pathways represent the suppression effect of exceptional situations on the acceptance of MP and MT. There is one direct and one indirect pathway. Figure 1 shows the causal structure of the model. The following paragraphs will discuss successively each of the particular paths in the model.

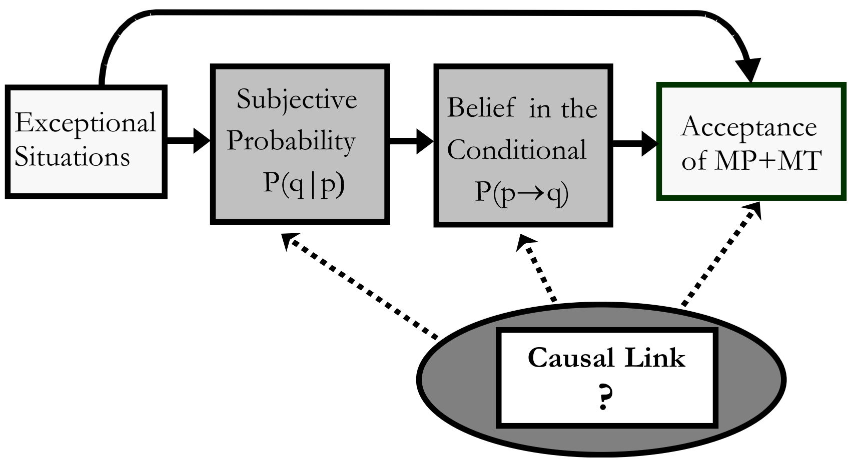


Figure 1: Integrated theoretical model of interpretation of and reasoning from causal conditionals: Availability of exceptional situations reduces the subjective conditional probability of consequent, given the antecedent; this in turn affects the degree of belief in the conditional. The acceptance of MP and MT inferences depends on the degree of belief in the conditional premise. In addition to the just specified indirect path there is a direct negative impact of the availability of exceptional situations on acceptance of MP and MT. The presence or absence of a causal link between antecedent and consequent might mediate these dependencies at various points.

2.1 Exceptional Situations → Subjective Probability P(q|p)

Because the conditional probability, *P(q|p)*, is a function of the probability of exceptional situations, it is assumed that people obtain an estimate of the probability of exceptional situations from their availability. The availability of exceptional situations increases the probability of exceptions to the conditional rule (i.e., cases of *p* together with *¬q*) and thus reduces the subjective conditional probability of the consequent, given the antecedent, *P(q|p)* (henceforth shortly conditional probability).

Thompson (1994) as well as Cummins (1995), seems to favor a similar approach in suggesting that disabling conditions exert their influence on MP and MT through reducing the perceived sufficiency of *p* for *q* (see also Fairley, Manktelow, & Over, 1999). Because there is no room for degrees in the logical concept of sufficiency, a better characterization for degrees of perceived sufficiency are varying subjective conditional probabilities, *P(q|p)*[3].

This link in the model is empirically supported through a negative correlation between availability of exceptional situations and the conditional probability of the consequent, given the antecedent that was found by Dieussaert, Schaeken, and d'Ydewalle (2002) in a study with more than 100 different conditional statements.

2.2 Subjective Probability P(q|p) → Belief in the Conditional

The model postulates that the subjective conditional probability, *P(q|p)*, corresponds to the overall degree of belief in the conditional statement, *P(p→q)*. The foundation of this inner part of the indirect pathway lies in probabilistic theories of the understanding of and the reasoning from conditionals. For example, following the suppositional theory Edgington (1991, 1995, 2003) proposes that the degree of belief a conditional deserves equals the conditional probability of the consequent *q* given the antecedent *p*. Probabilities representing degrees of confidence capture the intuition that many statements in everyday life are often not simply true or false, but can be trusted to a certain degree. For example, most gardeners count on the positive effects of fertilizers, even if occasionally a fertilizer is unsuccessful – for instance, because the plant suffers from varmints (exceptional situation).

In two experiments by Hadjichristidis and colleagues (2001), the estimated probability that a conditional is true (i.e., the degree of belief in a conditional, *P(p→q)*) correlated highly with the conditional probability of the consequent, given the antecedent. But is the degree of belief in a conditional really tuned to probabilities? This major prediction of the conditional probability approach (Edgington, 1991, 1995) has been tested recently in a series of experiments by Oberauer and Wilhelm (2003a) as well as by Evans et al. (2003). In these experiments participants received information about the frequencies of *pq*, *p¬q*, *¬pq*, and *¬p¬q*-cases. Participants estimated the likelihood of the conditional being true for one randomly selected case (Evans et al., 2003) or for a

---

[3] That is because sufficiency (of *p* for *q*) implies that *P(q|p)* = 1 and necessity (of *p* for *q*) implies *P(p|q)* = 1.

randomly drawn sample (Oberauer & Wilhelm, 2003a). In both series of experiments estimates were strongly affected by the ratio of *pq* to *p¬q*-cases that corresponds to the conditional probability, but also by the frequency of *pq*-cases. Edgington's conditional probability theory accounts for the former but not the latter. It is important to note that the ratio of *pq* to *p¬q*-cases affected the degree of belief in a conditional even when this was measured by simply asking participants whether the conditional was true or false (Experiments 2 and 4 by Oberauer & Wilhelm, 2003a). With this wording the possibility can be ruled out that the question itself triggers a probabilistic reading of the conditional.

Further support for the idea that everyday conditionals are understood as expressing probabilistic relations was presented by Byrne and Walsh (2002). They report that people who were confronted with a contradiction to a conditional inference they accepted beforehand by being informed that the conclusion turned out false in this instance, more frequently generated revisions of the interpretation of the conditional (e.g., "A's do not necessarily have B's" or "If *p* then probably *q*") than they "indicated disbelief, denial, rejection, doubt, or uncertainty about the conditional's truth" (Byrne & Walsh, 2002, p. 5).

A prominent probabilistic theory of conditional *reasoning* has been suggested by Oaksford, Chater, and Larkin (2000; Oaksford & Chater, 2001) following other researchers in this area (see for example Chan & Chua, 1994; George, 1997, 1999; Liu, Wu, & Lo, 1996; Stevenson & Over, 1995). According to Oaksford et al.'s theory "people endorse an inference in direct proportion to the conditional probability of the conclusion given the categorical premise" (Oaksford et al., 2000, p. 884). Therefore, the likelihood of acceptance of MP equals the conditional probability, $P(q|p)$, which in turn depends only on the probability of exceptions, *P(exceptions)*. Exceptions are true antecedent/false consequent-cases (i.e., *p¬q*). The likelihood that participants accept MT likewise depends on the probability of exceptions, but additionally on the prior probability of the antecedent, $P(p)$, and the consequent, $P(q)$. Consequently, Oaksford et al.'s theory predicts that the acceptance of MP and MT depends on exceptional situations and thus on the conditional probability, although this dependency might be stronger for MP than MT. According to this theory, the predicted connection between the probability of exceptions and the acceptance of MP and MT inferences is not explicitly mediated by the subjective probability that the conditional is true.[4]

2.3 Belief in the Conditional → Acceptance of MP and MT

The model assumes that the believability of the conditional premise measured as degree of belief affects the willingness to accept the valid inferences. There is a lot of evidence showing that MP and MT from believable conditional premises are accepted more frequently than from unbelievable ones (Byrne, 1989; George, 1995, 1997; Stevenson & Over, 1995; Stevenson & Over, 2001; Thompson, 1996; Torrens, Cramer, & Thompson, 1999). In most studies the same was true for the inferences AC and DA, although the

[4] Actually, in Oaksford et al.'s account the conditional as the major premise itself is completely irrelevant for the decision to accept or decline an inference. Liu (2003) has recently argued that people normatively should add a second computational step and conditionalize the output of the relevant conditional probability according to Oaksford et al. (2000) on the major conditional premise.

data are sparser and not entirely consistent. For a review on reasoning from uncertain conditionals see Politzer and Bourmaud (2002).

2.4 Direct Path from Exceptional Situations to Acceptance of MP and MT

Markovits and Barrouillet (2002; see also Barouillet & Lecas, 1998; Markovits, Fleury, Quinn, & Venet, 1998; Markovits & Quinn, 2002) proposed a mental model account of causal conditional reasoning that directly predicts that the availability of exceptional situations suppresses MP and MT. According to this approach and following the tradition of the mental model theory by Johnson-Laird (1983, 2001), people represent the premises in a set of mental models and reason from this set of mental models. Each model corresponds to a situation that fulfils the truth-conditions of the premises. Participants will accept a conclusion as deductively valid if and only if no counterexample can be found in which the premises are true and the conclusion is false. According to the theory of mental models of Johnson-Laird (1983), a counterexample is a mental model that is consistent with the premises, but contradicts the provisional conclusion. For example, the AC inference from the major premise "If a match is struck, it will light" and the minor premise "This match has been light" with the provisional conclusion "Therefore it has been struck", is contradicted by imaging a case in which a match lights, but has not been struck (e.g., because another burning match was held close to it). This mental model of q but $\neg p$ is analytically consistent with the premises, it neither contradicts the major nor the minor premise. This kind of counterexamples will be called *analytical counterexamples*. According to Johnson-Laird (1983) analytical counterexamples are essential for recognizing the so called fallacious inferences AC and DA as not valid. To the valid inferences MP and MT, there is no mental model that is consistent with the premises, but contradicts the putative conclusion, i.e., hence there are no analytical counterexamples to MP and MT. If the premises are true, a logically valid inference must be true too. Markovits and Barrouillet (2002) have extended the concept of analytical counterexamples to counterexamples that are logically independent from the premises. Confronted with the MP inference "A match is struck, therefore it lights", participants might imagine a situation in which a match was struck, but it did not light, because it was wet - based on general knowledge on matches. This is $p\neg q$-model contradicts the MT inference and constitutes what will be called an *empirical counterexample*. Although this situation is not compatible with the truth of the major premise, Markovits and Barrouillet (2002) assume that participants will reject the MP inference as well as the MT inference if the $p\neg q$-case becomes part of the mental model(s) of the premises.[5] Furthermore, the authors assume that relevant aspects of the context and content are automatically activated in the reasoner's memory whenever reading or hearing a conditional with realistic content. Because exceptional situations are highly relevant for a causal relation, they are automatically activated and if an exceptional situation is easily available, it is likely to be integrated into the current set of mental models. An exceptional situation is equivalent to

[5] The discrimination between analytical and empirical counterexamples was recently alleviated by introduction of the concept of "pragmatic modulation" into the mental model theory (Johnson-Laird & Byrne, 2002). By "pragmatic modulation" any relevant background information can become part of the mental model(s) of premises.

an empirical counterexample to the conclusions of MP and MT, because the corresponding model is *p¬q*. To keep the text simple and short, the term counterexample will refer to empirical counterexamples in the following - if not indicated otherwise.

Another revised mental model theory with an explicit probabilistic component was proposed by Schroyens and Schaeken (2003). These authors assume that participants perform a search for counterexamples (i.e., true minor premise/false putative conclusion cases) only if they are uncertain about the truth of the initial conclusion. If participants are able to retrieve an empirical or analytical counterexample during this search, they will reject the conclusion. This implies that MP as well as MT hinges on the likelihood of retrieving a counterexample to this inference (i.e., a *p¬q*-case). But the acceptance of MT does furthermore depend on the probability that participants consider the false-antecedent/false-consequent contingency, e.g. the likelihood of representation of the *¬p¬q*-case in participants' mental model(s). Schroyens and Schaeken (2003) conducted a meta-analysis of 65 studies: their mental model variant explained more variance in most studies than Oaksford et al.'s (2000) probabilistic theory.

Markovits and Barrouillet (2002) and Schroyens and Schaeken (2003) agree that whenever a reasoner becomes aware of an exceptional situation during the reasoning process, MP and MT will be rejected. Therefore, on average the probability of accepting MT or MP will depend on the probability of retrieving an exceptional situation. But it is important to emphasize that none of the process assumptions involve subjective probabilities. The link between accepting MP and MT and the availability of exceptional situations is not mediated through subjective probabilities in the reasoner's mind. Neither of the two modified mental model accounts embodies a representation of the conditional probability, $P(q|p)$, or of the probability of the conditional. Hence, there is no room within theses approaches for the idea that these believability estimates serve as mediators between retrieval of exceptional situations and the acceptance of an inference. To model and test this idea a *direct* inhibiting path from exceptional situations to the acceptance of MP and MT is included in the framework.

3 Experiments 1 to 3 (First Experimental Series)

A major goal of the first experimental series reported here was to test the integrated model empirically. To achieve this goal all variables in the model were assessed for a pool of conditional statements: availability of exceptional situations (Experiment 1 and 2), subjective conditional probability of consequent given the antecedent (Experiment 1), degree of belief in the conditional (Experiment 2), and acceptance of MP and MT in a conditional reasoning task (Experiment 3).

3.1 Research Questions

The model was developed on the basis of theoretical and empirical assumptions about causal conditionals and will be tested on causal conditionals at first. To examine furthermore the idea that the presence versus absence of a *semantic link* might play a role for the understanding and reasoning from the statement, the same conditionals were embedded in a cover story in which an explicitly mentioned causal link was either present or absent. Not only the presence versus absence of a causal was varied, but also the *structure* of the description of the causal relation. Forward causal statements (if cause, then effect) were compared with backward causal statements (if effect, then cause). Additionally, the impact of different instruction in reasoning tasks was examined by using two different instructions in the conditional reasoning task (Experiment 3): one with emphasis on the logical nature of the task and one with emphasis on an intuitive plausibility judgment. Theoretical considerations and predictions for each of the latter mentioned research goals will be presented in the following paragraphs.

3.1.1 The Role of the Causal Link: Causal and Noncausal Conditionals

In the introduction it has already been suggested that two conditional statements can differ in believability, even if they share a common conditional probability of the consequent given the antecedent. The connection between the propositions might modulate the believability over and above the conditional probability: According to the semantic link hypothesis conditionals with a meaningful relation between antecedent and consequent should have a higher believability.

A variety of researchers has suspected that arbitrary conditionals differ fundamentally from causal ones (Cheng & Holyoak, 1985; Markovits & Barrouillet, 2002; O'Brien, Costa, & Overton, 1986; Thompson & Mann, 1995; Peel, 1967; c.f. Bindra, Clarke, & Shultz, 1980). Nonetheless, empirical work concerning this question is rare: Few studies systematically compared noncausal and causal conditionals in the conditional inference task (for exceptions see Marcus & Rips, 1979; Valiña, Seoane, Ferraces, & Martin, 1999). Although the only way to control for a diversity of content and context effect is to compare the very same conditionals with and without a causal link, there is not one study yet that has used this method. The conditional statements that were used as items in the experiments later on had the same wording in both conditions (causal and arbitrary), and all were embedded in a pseudo-naturalistic context. Causal and arbitrary (henceforth noncausal) conditionals differed only in one respect: the presence versus absence of an explicitly mentioned causal relation between antecedent and consequent in the cover story. Similar types of material have been used by Waldmann (2001) in experiments on causal induction.

3.1.2 Forward and Backward Causal Links

Two structures of causal links were employed. The same causal relationship can be expressed as "If you fertilize a flower, then it will bloom" or as "If a flower blooms, then it has been fertilized". Following this, there were conditionals expressing a forward causal direction (if cause, then effect), and conditionals expressing a backward direction of causality (if effect, then cause)[6]. Forward causal conditionals are easier to understand than backward formulations (Bindra, Clarke, & Shultz, 1980; Traxler, Aked, Moxey, & Sanford, 1997). Mandel and Lehman (1998) asked participants for a definition of the term cause. Whenever participants used conditional statements to describe their concept of "cause", the cause took the role of the antecedent and the effect was represented as the

---

[6] Another way to describe forward and backward conditionals is to consider them as predictive and diagnostic conditionals (Waldmann, 2001; Waldmann & Holyoak, 1992).

consequent. Both findings support the intuition that forward wordings are the canonical way to state a causal relationship. Why should this be the case? What is the psychological difference between forward and backward formulations in causal conditionals? One of the reasons is probably the temporal order of events: Causes precede their effects. But additionally, it is suggested here that the way in which causes and effects are stored in human memory might play a role. This conjecture is motivated by theoretical and empirical work on conditional probability judgments by Gavanski and colleagues in category-feature relations (Gavanski & Hui, 1992; Hanita, Gavanski, & Fazio, 1997; McMullen, Fazio, & Gavanski, 1997; Sherman, McMullen, & Gavanski, 1992). They presented evidence that estimates of the structure P(category|feature), like "Of 100 randomly selected people who prefer blue rather than brown, how many are men?", were distorted towards the estimate of the inverted probability, P(feature|category), like "Of 100 randomly selected men, how many prefer blue rather than brown?". According to Gavanski and colleagues features are attached to categories in human memory, but there is no easy access to categories, given features as cues. Thus, when people are asked to estimate the probability of a category conditionalized on a feature, it often happens that they mentally reverse it, conditionalizing the feature on the category – because this is a task frequently and easily done. Tversky and Kahneman (1980) report asymmetries in conditional probability judgments in causal settings. For example, participants estimated the conditional probability that a girl has blue eyes, given that her mother has blue eyes, to be higher than the probability that a mother has blue eyes, given that her daughter has blue eyes: P(effect|cause) > P(cause|effect). Connecting these considerations yields the prediction that backward causal conditionals will sometimes be reversed to the forward form. This prediction is based on the assumption that cause-effect relations are organized in memory in a way similar to category-feature relations. Whenever a reversal happens, participants will represent the statement as "if q then p" instead of the original "if p then q". It is unclear what kind of consequences this will have on the estimates of the conditional probability and of the degree of belief. For the reasoning task, the predictions are clear-cut: Because the valid forms (MP and MT) turn into invalid forms (AC and DA), the reversal hypothesis predicts that acceptance rates of MP and MT decrease, whereas acceptance of AC and DA increase, for backward causal conditionals compared to forward causal conditionals.

Cummins (1995) and Thompson (1995) compared forward and backward causal conditionals in the conditional reasoning task. They were interested in whether the availability of disabling conditions suppressed MT and MP in forward, but AC and DA in backward causal conditionals. This was the case. Corresponding results were obtained for the availability of alternative causes which suppressed AC and DA in forward, but MT and MP in backward conditionals. If one rephrases this interaction independent of the causal link it means that the availability of exceptional situations suppresses MP and MT, and the availability of alternative situations, defined as cases in which the consequent q is true, but the antecedent p is false, suppresses AC and DA. This offers the possibility to transfer the model suggested here for MP and MT symmetrically to AC and DA inferences by substituting exceptional situations through alternative situations and exchanging p and q in all variables.

3.1.3 Deductive versus Inductive Instructions

Sometimes even minor differences between instructions can change the outcome of an experimental task dramatically. On the other hand, many effects are fairly robust to differences between instructions. The suppression of MT through easily available exceptional situations for example even persists with instructions that put heavy emphasis on the logical nature of the task (Vadeboncoeur & Markovits, 1999). Instructions can change the nature of the task and the strategies employed. Most research on causal conditionals underlying the model presented here did not instruct participants to reason deductively. It is therefore reasonable to assume that this is one reason for the strong effects of exceptional situations on the acceptance of inferences. What would happen if the instruction stressed that participants should decide only whether the inference follows with logical necessity? Rips (2001) has claimed that people have at least two distinct ways of evaluating arguments: deductively and in terms of their inductive strength. He compared inferences from highly plausible conditional premises ("If car X10 runs into a brick wall, it will stop") with inferences from very implausible premises ("If car X10 runs into a brick wall, it will speed up"). He found that people’s evaluations of the arguments depended much less on the plausibility of the conditional premise under deductive instructions than inductive instructions. Because availability of exceptional situations is probably closely linked to the plausibility of a conditional premise, Rips' variation of plausibility of the conditional premise can be regarded as a manipulation of the availability of exceptional situations. Hence, under deductive instruction it is expected that the effect of the availability of exceptional situations on the acceptance of inferences will be weaker under deductive than inductive instructions. Therefore, two instructions were used in Experiment 3: one with emphasis on logical necessity and one that put the task within a context of an everyday conversation in which participants had to judge the plausibility of the argument, corresponding to "deduction" versus "induction" instructions as used by Rips (2001).

3.2 Method

3.2.1 Participants

The experiments were run on the internet.[7] If one of the following criteria was met, the data were excluded from all analyses: missing answers to one or more items, acceptance of the question "Did you already participate in this experiment?", and an IP-number that had been used before in the same experiment. In Experiments 1 and 2 there remained 205 and 230 participants, respectively, and in Experiment 3 remained 1068 participants (499 in the induction group and 569 in the deduction group).

In Experiment 1 there were 102 men and 100 women, in Experiment 2, 116 men and 112 women, and in Experiment 3, 517 men and 542 women participated (the remainder did

[7] Experiments 1 to 3 were run in the Web-Lab at the University of Potsdam (http://w-lab.de) from December 2001 until December 2002. The experiments met most of the sixteen standards of internet-based experimenting that have recently been proposed by Reips (2002).

not respond to this question). The mean age was 28.6 (SD 9.5), 29.5 (SD 9.6) and 28.8 (SD 9.8) years for Experiment 1, 2, and 3, respectively (overall range 12 to 70).

3.2.2 Materials

The following conditionals were used:

1. *Allergic Disease*:
   a) If a dog suffers from Midosis, then one finds Xathylen in its blood.
   b) If one finds Xathylen in a dog's blood, then it suffers from Midosis.
2. *Tropical Plant*:
   a) If a Pherotelia blooms, then there are blue point beetles on it.
   b) If there are blue point beetles on a Pherotelia, then it blooms.
3. *Mechanical Object of Art*:
   a) If the light is on, then the song is playing.
   b) If the song is playing, then the light is on.
4. *DNA-Mutation*:
   a) If one finds Natrolsan in a rabbit's body, then one finds the DNA-mutation.
   b) If one finds the DNA-mutation in a rabbit, then one finds Natrolsan in its body.
5. *Tribal Behavior*:
   a) If a man smokes Zenobia herbs, then he suffers from hair loss.
   b) If a man suffers from hair loss, then he smokes Zenobia herbs.
6. *Computer Virus*:
   a) If the PC's date is 12-24-2001, then the internet browser's starting page is www.joke.com.
   b) If the internet browser's starting page is www.joke.com, then the PC's date is 12-24-2001.
7. *Social Relationship*:
   a) If Katrin inquires with importunity, then Michael retreats into silence.
   b) If Michael retreats into silence, then Katrin inquires with importunity.
8. *Outer Space Physics*:
   a) If a probe is high on Philoben gas, then it has at least 22 degree centigrade.
   b) If a probe has at least 22 degree centigrade, then it is high on Philoben gas.
9. *Alarm Equipment*:
   a) If the floodlight is on, then the siren is howling.
   b) If the siren is howling, then the floodlight is on.

As the list shows, the nine conditional statements came from very different semantic domains. With a large variety of scenarios it was hoped to provide a fair test for a general model. To control for order effects, both possible orders of antecedent and consequent term were used for each conditional resulting in versions a) and b). Every problem (i.e., item) consisted of a cover story plus a conditional statement. For each of the nine conditionals three slightly different cover stories were created to manipulate the causal structure of the problems. The basic paradigm and one conditional were adapted from

Waldmann (2001). There were two sorts of causal stories (implying *forward* and *backward* direction of causality) and a neutral cover story introducing the conditional as an *arbitrary* connection between two elementary propositions (noncausal problems). The word "cause" was not used in any of the cover stories. Here is an example for a causal forward problem (if cause, then effect):

> *A laboratory in Australia has recently discovered a new allergic disease in dogs. The new disease has been named Midosis. Since that time the researchers have discovered many characteristics of Midosis. Among other things, the scientists have detected that the disease makes an affected dog's blood produce the formerly unknown substance Xathylen.*
>
> *Sara is a practicing veterinarian. She assumes that it generally holds that:*
> *'If a dog suffers from Midosis,*
> *then one finds Xathylen in its blood'*

The second causal cover story for the same conditional was almost identical to the first, but instead of Midosis causing Xathylen to be produced, the story explained that Xathylen is the substance that leads to the symptoms of the allergic disease Midosis. Because the causal roles of the two events are reversed, the conditional now had the form "if effect, then cause" instead of "if cause, then effect". This yields a causal backward conditional (if effect, then cause):

> *A laboratory in Australia has recently discovered a new allergic disease in dogs. The new disease has been named Midosis. Since that time the researchers have discovered many characteristics of Midosis. In the blood of the affected dogs, a formerly unknown substance called Xathylen is produced. This substance leads to the multiple symptoms of Midosis.*
>
> *Sara is a practicing veterinarian. She assumes that it generally holds that:*
> *'If a dog suffers from Midosis,*
> *then one finds Xathylen in its blood'*

The neutral cover story for the noncausal conditionals was similar to the causal stories. The major difference was that the story did not make any explicit causal connection between Midosis and Xathylen. The link between antecedent and consequent was purely arbitrary, because no meaningful connection was suggested through the cover story:

> *The major fields of research of a laboratory for veterinary medicine in Australia are the research on physiology (e.g., the compounds of blood in different animals) and allergic diseases. The laboratory has recently discovered a new allergic disease in dogs called Midosis. A different department has detected a formerly unknown substance in the blood of cats and has named it Xathylen. During the last weeks the scientists explored whether Xathylen is found in dogs' blood as well.*
>
> *Sara is a practicing veterinarian. She assumes that it generally holds that:*
> *'If a dog suffers from Midosis,*
> *then one finds Xathylen in its blood'*

The combination of nine conditionals with three cover stories resulted in a total of 27 problems (different combinations of conditional statement plus cover story). As already mentioned, both possible orders of antecedent and consequent term were used for each conditional (versions a and b, e.g. "If Midosis, then Xathylen" as well as "If Xathylen, then Midosis"). This was done to control for order effects. Therefore the item pool consisted of 54 problems. All materials were written in German, and can be found completely in Appendix A.

3.2.3 Procedure

Experiment 1 and 2 were identical except for the first question: Participants in Experiment 1 estimated the conditional probability of the consequent, given the antecedent, that is, $P(q|p)$, and participants in Experiment 2 estimated the probability of the conditional being true, that is, $P(p \rightarrow q)$.

In Experiment 1 and 2 each participant received six problems (two forward, two backward, two noncausal problems), each with a different thematic content (1-9, see above), in random order. Each problem was presented on a separate screen. All problems were followed by four questions. Following Sara's hypothesis: "If a dog...", participants were asked in Experiment 1:

1. *A dog is randomly selected from the lab's kennel. It turns out that this dog suffers from Midosis. What do you think how likely it is that one finds Xathylen in its blood?*
   Please enter a number between **0** *(***totally impossible***) and* **100** *(***absolutely certain***).*

or in Experiment 2

1. *What do you think how likely it is that Sara's statement holds true?*
   Please enter a number between **0** *(***totally impossible***) and* **100** *(***absolutely certain***).*

The rest of the questions was identical in Experiments 1 and 2. The second question was a rating of the perceived causal strength[8], and the last two questions assessed the availability of exceptional situations (cases of *p* but *not-q,* shortly $p\neg q$ cases) and of alternative situations (cases of *q* but *not-p,* shortly $q\neg p$ cases):[9]

2. *Do you think that Sara's statement describes a causal relationship?*
   - *No, the statement does not describe a causal relationship.*

   If you think it does, please estimate the **strength** *of this causal relationship:*
   (a 5-point-rating scale from very weak to very strong causal relationship was provided here)

---

[8] Some theories assume that causal strength is highly relevant for reasoning in causal domains (e.g. Quinn & Markovits, 1998).

[9] Alternative situations are circumstances that lead to the consequent *q* in absence of the antecedent *p*. For forward causal conditionals this corresponds to alternative causes and for backward causal conditionals to disabling conditions.

3. *For the next question please assume that Sara's statement is true. Can you imagine conditions under which the following instance is possible? A dog suffers from Midosis, but no Xathylen is found in its blood.*

4. *For the next question please assume that Sara's statement is true. Can you imagine conditions under which the following instance is possible? Xathylen is found in a dog's blood, but it does not suffer from Midosis.*

For the questions 3 and 4 participants indicated how many conditions they could think of on a five-point rating scale from "*very few*" to "*very many*", with an additional option provided "*No, I cannot imagine such a situation*". To ensure that participants carefully considered the conditions, they were furthermore asked to describe one situation in short keywords. The order of presentation of the questions concerning "*p* but *not-q*" (exceptional situations) and "*q* but *not-p*" (alternative situations) was chosen randomly for each item.

For half of the noncausal problems a control question was included that checked for the possibility that participants spontaneously adopted a causal interpretation of the conditionals - even with neutral cover stories. The control question was inserted in these problems instead of the two questions concerning the availability of exceptional and alternative situations. For these items participants were asked to which statement they would consent most:

1. *Midosis causes Xathylen to be in the blood.*
2. *Xathylen in the blood causes Midosis.*
3. *Neither one causes the other.*
4. *There is a third common cause to both.*

The order of presentation of the first two options ("*p* causes *q*" and "*q* causes *p*") was randomized for each problem.

In Experiment 3 participants solved conditional reasoning problems under two instructions. The relevant part of the instructions read as:

The tasks in this experiment will have the following form:

***Assumption**:*	*"If Anna travels to Italy, then she takes the train."*
***Fact**:*	*Anna travels to Italy*
***Conclusion**:*	*Anna takes the train*

Please imagine the tasks to be part of a conversation or a discussion. Your discourse partner draws the conclusion from the assumption and the fact.

(For the deduction group only:)
Please examine whether the argumentation is logically coherent. ***Don't*** *judge whether the assumption and the fact may be the case, but instead whether the conclusion follows with* ***logical necessity*** *from the assumption and the fact. That is to say that if the assumption and the fact are true, the conclusion cannot be false.*

(For the induction group only:)
Please examine whether you find the argumentation ***plausible****. Would you follow the argumentation of your discourse partner and accept the conclusion?*

It was stressed in both groups that the relation in an if-sentence is not necessarily reversible. Each participant then received three problems (one forward, one backward, one noncausal) from different semantic domains (1-9, see above), presented in random order, each on a separate screen. Each problem was followed by four inferences: MP, MT, AC, and DA. Here is an example for MP:

- Cover story -

Sara is a practicing veterinarian. She assumes that it generally holds that:
"If a dog suffers from Midosis,
then one finds Xathylen in its blood"

Fact*: This dog suffers from Midosis.*
Sara concludes*: Xathylen is found in its blood.*

1. *Is Sara's Conclusion logically valid? (*deduction group)
   YES *NO*

or

1. *Do you think that Sara's conclusion is plausible?* (induction group)
   YES *NO*

2. *How confident are you in your decision?*
   (a 4-point-rating scale from *uncertain* to *certain* was provided here)[10]

To control for order effects two different orders of inferences were used: MP-AC-MT-DA and DA-MT-AC-MP. Within each participant order of presentation of inferences was held constant. Complete instructions and examples of original items (both in German) can be found in Appendix C.

[10] Confidence ratings did not yield any interesting results and will therefore not be discussed.

3.3 Results

3.3.1 Manipulation Check and First-Order Correlations

To ensure that the conditionals in noncausal problems were not spontaneously understood as expressing causal relations a manipulation check was performed. To make furthermore sure that the causal conditionals are comparable to natural causal conditionals that have been used before, e.g., by Cummins (1995), correlations between the following variables were computed: availability of exceptional situations, availability of alternative situations, acceptance of MP, MT, AC, and DA. The pattern of correlations for the causal conditionals was compared to the pattern expected from the literature.

A control question that examined whether participants spontaneously attributed causality to the noncausal conditionals was included in about half of the items in Experiments 1 and 2. Table 1 shows two important findings: Most people in most semantic domains did not spontaneously adopt a causal interpretation of the noncausal conditionals. This justifies the conclusion that the manipulation of the presence versus absence of a causal link can be regarded as successful, although the frequency of the neutral answer "Neither one causes the other" varied largely between semantic domains (from 14.6 % for alarm equipment to 84.0 % for allergic disease). It is furthermore important to note that the differences in answering patterns between the semantic domains demonstrate that heterogeneous material was created – as intended. There was a trend to take on the causal view that "*p* causes *q*" more often than "*q* causes *p*", but this difference did not reach the conventional level of alpha error, X^2 (1) = 3.0, $p = 0.083$.

To make sure that the causal conditionals used were similar to natural causal conditionals, correlations over the causal items ($n = 36$) were examined. There was a strong negative correlation between availability of exceptional situations and the acceptance of MP ($r = -0.81$, $p < 0.001$) and MT ($r = -0.76$, $p < 0.001$), but not with the AC or DA inference form. The reverse was true for alternative situations: Availability of alternative situations correlated negatively with AC ($r = -0.80$, $p < 0.001$) and with DA ($r = -0.75$, $p < 0.001$), but not with MP or MT. The pattern of correlations clearly indicates that the causal items are similar to natural causal conditionals.

Table 1: Answers to the Control Question for Spontaneous Attribution of Causality (Noncausal Conditionals Only)

	Relative frequency in percent Overall and per scenario									
	Overall	1	2	3	4	5	6	7	8	9
p causes *q*.	12.0	4.0	10.5	6.7	16.7	10.8	10.3	27.1	14.6	7.3
q causes *p*.	8.0	4.0	5.3	2.2	10.0	8.1	0	18.8	10.4	12.2
Neither one.	52.7	84.0	55.3	53.3	60.0	75.7	48.7	29.2	54.2	14.6
Third common cause.	27.4	8.0	28.9	37.8	13.3	5.4	41.0	25.0	20.8	65.9

Note. Neither one = Neither one causes the other. Third common cause = There is a third common cause to both. 1 = Allergic Disease, 2 = Tropical Plant, 3 = Mechanical Object of Art, 4 = DNA-Mutation, 5 = Tribal Behavior, 6 = Computer Virus, 7 = Social Relationship, 8 = Outer Space Physics, 9 = Alarm Equipment. Overall N = 376.

Among the inference forms MP correlated with MT ($r = 0.85, p < 0.001$) and AC with DA ($r = 0.76, p < 0.001$). Additionally, MT correlated with DA ($r = 0.48, p < 0.01$). MT and DA both include negated propositions. Schroyens and Schaeken (2003) suggested a mental model variant in which the acceptance of MP and MT depends on the likelihood that a reasoner takes into account an exceptional situation ($p\neg q$) as a possibility, while AC and DA both hinge on the likelihood that an alternative situation ($\neg pq$) is considered. Furthermore, MT and DA are accepted only if a reasoner incorporates the false-antecedent/false-consequent contingency, that is the mental model of $\neg p \neg q$. Therefore, according to the model of Schroyens and Schaeken (2003), the inferences share parameters pairwise: MP with MT, AC with DA, and MT with DA. The pattern of correlations between the inferences matches nicely these shared parameters of the inference forms. For a complete overview over the intercorrelations between experimental variables please see Appendix B.

3.3.2 Testing the Integrated Model with a Path Analysis

The integrated model of understanding and reasoning from conditionals introduced above in chapter 2 (pp. 3-8) can be translated completely into a causal path model. This has been done in Figure 2. To avoid transfer effects that would artificially increase correlations between variables, different groups of participants estimated the conditional probability, the degree of belief in a conditional, and the acceptability of the inferences under deductive and under inductive instruction. Therefore, it was not possible to conduct a path analysis over participants. Instead, correlations across problems were computed to conduct a path analysis across problems (i.e., items). Mean estimates of all relevant variables in the model were computed for each of the 54 problems: mean values of the conditional probability, the probability of the truth of the conditional, and the acceptance rates of the MP and MT inference. One consequence of this procedure is a relatively small sample size ($N = 54$). Interpreting the path analysis appears nonetheless justified, since the small sample size was probably compensated by highly reliable estimates of correlations: Each data point rests on aggregated responses from 20-30 participants.[11]

The path model depicted in Figure 2 was fitted to the overall data, allowing different path coefficients for the causal statements (taking forward and backward causality together), and the noncausal statements. This approach yields an overall model fit, and additionally allows looking at two distinct sub-models: one sub-model for the causal conditionals and one sub-model for the noncausal conditionals. Figure 2 displays the standardized path coefficients and the explained variance (R^2-values) for the dependent variables for the causal sub-model. The overall model fit was satisfactory, as can be seen from the fit indices listed in Table 2.

The two sub-models (causal conditionals alone and noncausal conditionals alone) can be compared with respect to the model fits (see Table 2) as well as with respect to the size and the significance of the path coefficients, which are listed in Table 3. The sub-

[11] With one exception: Because half of the noncausal items had a control question instead of the assessment of the availability of exceptional situations, there were 40-60 answers per item for the availability of exceptional situations for the *causal* conditionals.

model for the causal conditionals alone had a good fit. All path coefficients were significantly larger than zero except one that was only marginally significant. For the sub-model with the noncausal conditionals alone, the fit was not convincing. The paths connecting availability of exceptional situations, subjective conditional probability, and the degree of belief in the conditional were almost the only paths that obtained significant coefficients. These paths are the first steps in the *indirect pathway*. These paths can furthermore be constrained to the same value in both sub-models (causal and noncausal) without serious decline in the fit indices (this is constraint 1 in Table 2).

Table 2: Measures for Goodness of Fit[12] for Path Models with Different Constraints

		X^2	RMSEA	CFI
No constraints:	overall	29.7 (*df* = 22, *p* = 0.126)	0.082 (0-0.151)	0.970
	causal alone	10.6 (*df* = 11, *p* = 0.477)	0.000 (0-0.172)	1.000
	noncausal alone	18.8 (*df* = 11, p = 0.065)	0.205 (0-0.358)	0.877
Constraint 1 (between models):	overall	30.2 (*df* = 24, *p* = 0.178)	0.071 (0-0.140)	0.975
Constraint 2 (within models):	overall	32.5 (*df* = 26, *p* = 0.177)	0.069 (0-0.136)	0.974
	causal alone	11.8 (*df* = 13, *p* = 0.544)	0.000 (0-0.155)	1.000
	noncausal alone	20.4 (*df* = 13, *p* = 0.086)	0.183 (0-0.328)	0.885
Constraint 1+2 simultaneously:	overall	33.1 (*df* = 28, *p* = 0.232)	0.059 (0-0.127)	0.980

Note. Constraint 1 = paths "exceptional situations → subjective probability" and "subjective probability → belief in the conditional" were fixed to be equal in the causal and noncausal sub-model. Constraint 2 = within the sub-models the right side (deduction group) and the left side (induction group) were fixed to the same value. The "causal alone" and "noncausal alone"-models are sub-models of the overall model.

The results of the structural equation model show no evidence of differences between the instruction groups in the reasoning task. The left side of the path diagram, with inferences from the deduction group as dependent variables, does not differ from the right side, where the dependent variables are inferences drawn in the induction group. When the paths for the deduction group (left side) were fixed to be equal to the corresponding paths for the induction group (right side), the overall model fit decreased only negligibly (this is constraint 2 in Table 3). This constraint applied within, not between, the sub-models.

When both stated constraints (first steps of the indirect path fixed to the same value in the causal and the noncausal sub-model, and values for the deduction group fixed to the same value as the induction groups) were applied simultaneously, the pattern of path coefficients is clear-cut: For the causal conditionals all path coefficients received significant weights. For the noncausal conditionals all path coefficients received significant weights except for the paths from the belief in the conditional to the latent variables "Logic" and "Plausibility" that correspond to the proportions of participants that accepted the MP and MT form in the deduction and induction group, respectively.

---

[12] RMSEA (Root Mean Square Error of Approximation) indices are reported with upper and lower bound in parentheses. A good fit is indicated by an RMSEA index smaller than 0.05, an acceptable fit by a RMSEA index smaller than 0.08. A RMSEA larger than 0.10 displays a poor fit.
A CFI (Comparative Fit Index) between 0.90 and 0.95 indicates an acceptable fit, values above 0.95 a good fit and below 0.90 a poor fit.

Remarkably, the depicted patterns of results did not change if the roles of subjective conditional probability and degree of belief in the conditional were swapped within the framework. This is compatible with the assumption that the two variables might be indicators of the same construct.

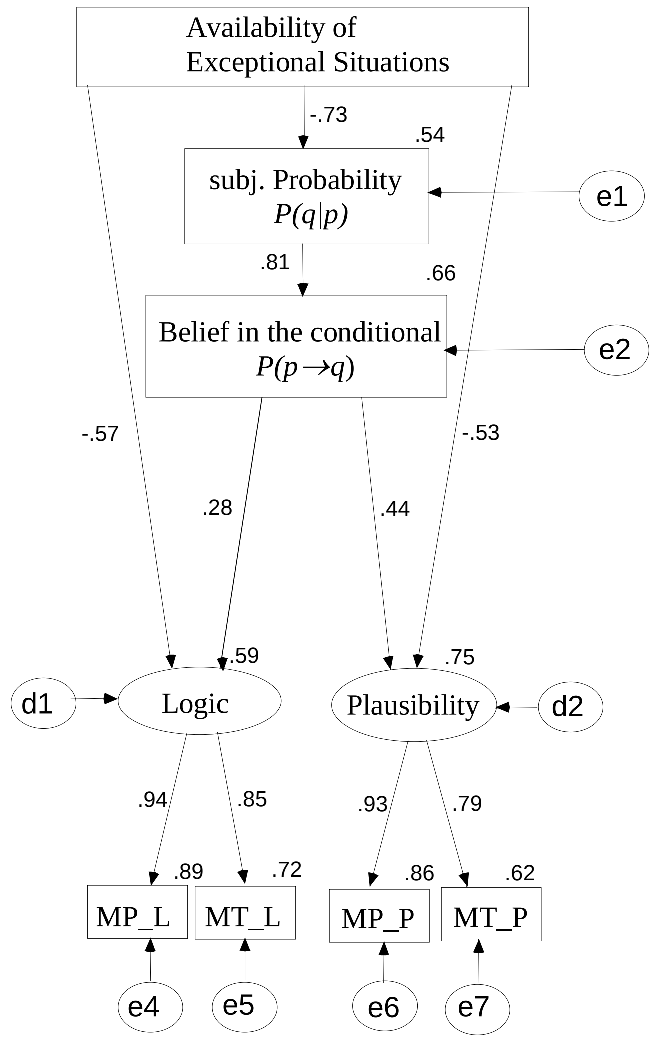


Figure 2: Structural equation model and standardized results for the causal conditionals. "Logic" is a latent variable defined by the proportion of participants in the group instructed to reason logically who accepted MP and MT (deduction group). "Plausibility" is the corresponding latent variable for the group instructed to assess the argument's plausibility (induction group).

Table 3: Standardized Path Coefficients and *p*-Values for Causal and Noncausal Items

		Causal items (n = 36)		Noncausal items(n = 18)	
		Estimate	p	Estimate	p
Path					
Exceptional Situations	- Subjective Probability	-0.73*	0.00	-0.54*	0.01
Subjective Probability	- Belief	0.81*	0.00	0.85*	0.00
Exceptional Situations	- Logic (Deduction)	-0.57*	0.00	-0.08	0.68
Belief	- Logic (Deduction)	0.28	0.07	0.30	0.13
Exceptional Situations	- Plausibility (Induction)	-0.53*	0.00	-0.76*	0.00
Belief	- Plausibility (Induction)	0.44*	0.00	-0.01	0.95
Logic (Deduction)	- MP_L	0.94	(fixed)	1.19	(fixed)
Logic (Deduction)	- MT_L	0.85*	0.00	0.66	0.16
Plausibility (Induction)	- MP_P	0.93	(fixed)	0.69	(fixed)
Plausibility (Induction)	- MT_P	0.79*	0.00	0.89*	0.01

Note. Estimate = Estimate of the standardized path coefficient. Belief = Belief in the Conditional. For "Logic" and "Plausibility" see caption of Figure 2. No constraint applied to the model. *p < 0.05.

3.3.3 Causal Structure

The effects of the causal structure of the conditionals (causal forward vs. causal backward vs. noncausal items) were investigated in two separate analyses: one analysis with the two believability estimates from Experiment 1 and 2, and one analysis with the reasoning data from Experiment 3.

The data from Experiment 1 and Experiment 2 were submitted to an analysis of variance with type of conditional (causal forward vs. causal backward vs. noncausal problems) as within participants factor and group (estimates of conditional probability vs. estimates of degree belief in a conditional) as between factor. Estimates of the conditional probability (mean 74.4) were higher than estimates of the degree of belief in the conditional (mean 68.7), as indicated by a main effect of group, $F_{(1, 433)} = 68.2$, $p < 0.001$, $Eta^2 = 0.136$. Examination of Table 4 reveals that forward conditionals received higher estimates than backward conditionals, $F_{(1, 433)} = 36.9$, p < 0.001, $Eta^2 = 0.079$, and causal conditionals (forward and backward) received higher estimates than noncausal ones, $F_{(1, 433)} = 306.8$, $p < 0.001$ $Eta^2 = 0.415$, which yielded two significant post hoc contrasts of the main effect of type of conditional, $F_{(2, 432)} = 176.8$, $p < 0.001$, $Eta^2 = 0.290$. Estimates of conditional probability and of belief in a conditional were closer to each other for the two types of causal conditionals than they were for the noncausal conditionals, as indicated by an interaction between group and type of conditional, $F_{(2, 432)} = 9.1$, $p < 0.001$, $Eta^2 = 0.020$.

Table 4: Mean Estimates for the Three Types of Conditionals in Experiment 1 and 2 (Standard Deviations in Parentheses)

	Causal forward	Causal backward	Noncausal
Subjective probability	77.6 (19.6)	71.3 (22.5)	57.2 (24.2)
Degree of belief	71.9 (20.8)	61.0 (23.5)	39.3 (23.5)

Note. Entries are mean estimates of subjective conditional probability of q, given p, (subjective probability) and mean estimates of subjective degree of belief in the conditional (degree of belief), respectively, both given on a scale from 0 to 100.

Table 5: Percentages of YES-Answers to Inferences for the Three Types of Conditionals and Two Instruction Groups

	Causal forward	Causal backward	Noncausal
MP	84.7 / 77.9	77.1 / 66.6	72.5 / 62.1
MT	57.7 / 64.8	54.2 / 51.9	48.2 / 43.9
AC	47.8 / 54.8	51.1 / 51.3	36.8 / 36.3
DA	46.8 / 54.9	52.8 / 60.3	34.0 / 37.0

Note. The rows contain the percentage of acceptance of the four inference forms for each of the two instruction groups: First entry is the value for the deduction group, second entry the one for the induction group.

Table 5 gives an overview over the means of the *acceptance data* in the reasoning task (Experiment 3). The data were submitted to an analysis of variance with inference form (MP+MT vs. AC+DA) and type of conditional (causal forward vs. causal backward vs. noncausal) as within participants factors and instruction (deduction vs. induction) as between participants factor. The analysis revealed that the valid forms MP and MT were accepted more often than the invalid forms AC and DA as indicated by a main effect of inference form, $F_{(1, 1011)} = 201.7$, $p < 0.001$, $Eta^2 = 0.166$ (mean acceptance rates were 0.63 and 0.47 for MP+MT and AC+DA, respectively). The main effect of type of conditional was significant as well, $F_{(2, 1010)} = 83.5$, $p < 0.001$, $Eta^2 = 0.076$: Overall, inferences from forward causal conditionals and backward causal conditionals (means were 0.61 and 0.58, respectively) were accepted more often than inferences from noncausal conditionals (mean 0.46). Furthermore, inference form interacted with type of conditional, $F_{(2, 1010)} = 25.8$, $p < 0.001$, $Eta^2 = 0.025$. This interaction is depicted in Figure 3. The figure shows that the valid forms (MP+MT) were accepted more often with forward than with backward conditionals, whereas the invalid forms (AC+DA) were accepted more often with backward than with forward conditionals. Thus, there were more logically correct answers to the forward than to the backward causal conditionals – this was true for the valid as well as the invalid forms.

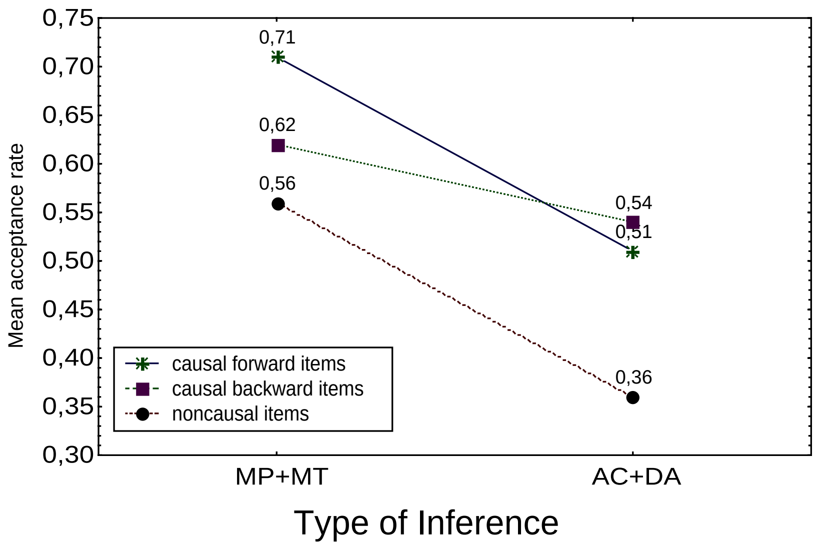


Figure 3: Interaction between type of conditional and type of inference

There was no main effect of instruction group ($F < 0.1$), but instruction interacted in an unexpected way with type of conditional[13], $F_{(2, 1010)} = 4.8$, $p = 0.008$, $Eta^2 = 0.005$, and with inference form, $F_{(1, 1011)} = 13.7$, $p < 0.001$, $Eta^2 = 0.013$. Detailed post hoc analyses for each inference showed that the differences between induction and deduction group were restricted to the inferences MP and AC. Participants in the deduction group answered more frequently according to the normative standard of the material implication than the induction group: they accepted MP more frequently (mean acceptance rates were 0.78 and 0.69 for deduction and induction group, respectively; Wald $X^2(1) = 33.6$ and $p < 0.001$) and rejected AC more frequently (mean acceptance rates were 0.45 and 0.51 for deduction and induction group respectively; Wald $X^2(1) = 12.2$ and $p < 0.001$). But there was no difference between the two instruction groups with respect to the acceptance rates of MT and DA. These results indicate that participants were sensitive to the difference between the two instructions, although the effect of the instruction was rather small and restricted to the (easy) inferences MP and AC.

3.3.4 *Reasoning Patterns*

Every participant solved all four inferences for each problem he/she received. Therefore, it is possible to examine whether there are systematic patterns of responses to all four inferences to be observed in the answers. A biconditional pattern for example would be observed if a participant accepted all inferences to a given problem. The analysis of patterns is especially interesting for the two sorts of causal conditionals, since the *reversal hypothesis* assumes that in backward causal conditionals participants sometimes reason from the reverse conditional "if *q* then *p*" instead of "if *p* then *q*". This implies that valid forms turn into invalid forms and vice versa, this should lead to a very unusual pattern: MP and MT will be rejected, while AC and DA are accepted (reverse conditional pattern).

An alternative explanation for differences between forward and backward causal conditionals is that backward causal items favor a biconditional reading. If the canonical way to phrase a causal relationship in a conditional statement is "if cause, then effect", then the content of the cover story for backward problems suggests the statement "if *q* then *p*". However, the conditional statement that follows the cover story asserts explicitly "if *p* then *q*". Reasoners might represent both conditionals which would result in an increase of biconditional readings of backward items. This idea will be called *equivalence hypothesis*.

---

[13] To explore the interaction of type of conditional with instruction in detail, we conducted a separate analysis for each of the four inferences with a generalized Chi^2-test (Wald's test) with type of conditional and group as factors. These post hoc tests revealed that the observed interaction between type of conditional and group in the overall test resulted solely from the MT inference (Wald $X^2(2) = 8.0$, $p < 0.05$): The induction group accepted more MT inferences than did the deduction group, but only in the causal forward condition (acceptance rates were 0.65 and 0.58 for the induction group and the deduction group, respectively). For the other types of conditionals (causal backward and noncausal) there was a trend in the opposite direction: for these sorts of items participants in the deduction group accepted the MT inference more frequently (acceptance rates were 0.48 and 0.51 for the induction group and the deduction group, respectively). This latter trend would be expected if participants followed the instruction. Somehow the forward causal conditionals seem to favor the MT inference for the induction group. I don't know of any theory that can account for this effect and see no way to decide at this moment whether it is due to the specific materials used, an accidental effect, or of theoretical importance.

By taking a look at the pattern of responses, reversal and equivalence hypothesis can be distinguished from each other. After items with missing answers were excluded, 3115 problems remained for which a classification into answering patterns was possible. The equivalence hypothesis predicts an increase in biconditional answers for backward items relative to forward causal items. The reversal hypothesis predicts that the reverse conditional pattern is more prevalent in backward items than in forward items. The data for each pattern were submitted to a generalized Chi^2-analysis (Wald's test) with type of conditional (causal forward vs. causal backward vs. noncausal) and instruction (deduction vs. induction) as factors; post hoc tests were conducted if necessary.

Overall, a biconditional pattern was observed in about a third of the cases. This pattern was more frequent for noncausal than for causal statements and more frequent for the induction group (39.3 %) than for the deduction group (31.1 %), Wald $X^2(1) = 22.6$, $p < 0.001$. But biconditional answers were by no means more frequent for backward than for forward items – this finding clearly contradicts the equivalence hypothesis. There was even a trend in the opposite direction, as apparent from Table 6.

The conditional pattern was more prevalent in the deduction group (19.9 %) than in the induction group (15.2 %; Wald $X^2(1) = 11.7$, $p < 0.001$). As mentioned above, at least some participants were sensitive to the instructions, and answered more in accordance with the normative standard in the deduction group than in the induction group. Furthermore the conditional pattern was more prevalent for the forward than for the backward causal items, Wald $X^2(1) = 7.2$ and $p = 0.007$.

The reverse conditional pattern was expected to be very rare, mainly because it entails the rejection of the almost universally accepted MP inference. Overall five percent of answers were classified as reverse conditional pattern. That corresponds to 161 occurrences. The frequency of the reverse conditional pattern did not differ between the instruction groups. But the reverse conditional pattern was almost twice as common in the backward conditionals relative to the forward condition, this difference is highly significant, Wald $X^2(1)= 12.1$, $p < 0.01$. This finding provides strong evidence for the reversal hypothesis – particularly because it is complemented by the reciprocal effect in the conditional patterns. While the conditional pattern was more frequent for forward than backward items, the reverse conditional pattern was more frequent for the backward items.

Table 6: Percentage of Reasoning Patterns for the Three Types of Problems

	Causal forward (n = 1039)	Causal backward (n = 1039)	Noncausal (n = 1037)
Biconditional	34.2	32.0	39.3
Conditional	19.6	15.2	17.9
Reverse conditional	4.4	8.2	3.1
Other	41.8	44.6	39.8

Note. Biconditional = MP+MT+AC+DA accepted. Conditional = MP+MT accepted, AC+DA rejected. Reverse conditional = MP+MT rejected, AC+DA accepted. All other patterns were classified as "Other". Data are aggregated over instruction groups.

3.4 Discussion

Key features of probabilistic, semantic, and mental model approaches of understanding and reasoning from causal conditionals have been integrated into one model. This model was tested through a path analysis with the data from Experiments 1 to 3. The analysis supported the model for causal but not for noncausal conditionals. In the model the suppression effect that the availability of exceptional situations exerts on MP and MT is modeled through two pathways that represent rival theoretical approaches. The indirect pathway via estimates of believability of the conditional (conditional probability and degree of belief in the conditional) is connected to probabilistic theories. The direct pathway is based on mental model variants by Markovits and Barrouillet (2002) as well as Schroyens and Schaeken (2003). By estimating the weights that the pathways receive within a causal path analysis, the relative merits of the two groups of theories can be assessed. Surprisingly, both pathways received significant weights for the sub-model of the causal conditionals alone. If the indirect pathway as well as the direct pathway from exceptional situations to the acceptance of MP and MT receives significant weights, one has to deduce inevitably that in the domain of causal conditional reasoning the two groups of theories are not mutually exclusive, but instead both have their merits. This finding raises the question how these quite different theories can be reconciled. At this point, only speculations can be offered on this question.

One possible explanation is based on the differences between the tasks: Obviously, reasoning from a conditional statement is different from evaluating a conditional's believability or its truth. In reasoning, a minor premise comes into play as an additional element. On the ground of recent findings of Markovits and Potvin (2001) the assumption is warranted that reading and representing a minor premise (re)activates the whole knowledge structure in semantic memory associated with the conditional statement. A highly relevant part of this knowledge structure are disabling conditions and alternative causes (i.e., exceptional situations) and this additional activation through the minor premise could explain why the availability of exceptional situations has an effect on reasoning that is independent from its effect on the believability of the conditional statement.

Alternatively, it should be considered that the two pathways might represent different mechanisms. Verschueren, Schaeken, and d'Ydewalle (2003) successfully specified and tested a dual-process theory (Evans and Over, 1996; Schroyens, Schaeken, & Handley, 2003) for causal conditional reasoning. One of the reasoning mechanisms is assumed as a *heuristic* mechanism that uses probability information to reach a degree of confidence in a conclusion, concretely this information is $P(q|p)$ for MP and MT, and $P(p|q)$ for AC and DA. The relevant probability is retrieved from long-term memory where it is estimated from a range of relevant situations. This probability estimation process is preconscious, fast, and computationally powerful. The second postulated mechanism is specified with a mental model approach and called *analytical*. The analytical reasoning mechanism searches for counterexamples in memory: An initial conclusion is rejected if a counterexample is found in memory. Verschueren et al. (2003) assume that both processes start simultaneously and operate over a certain time period, within a critical time frame. The output of the heuristic system is quickly available, but can subsequently be overridden by the output of the analytical system. After the critical time frame ends, the conclusion is rejected if a counterexample was found in time by the analytical mechanism. If not, the reasoner falls back on the degree of belief in the inference given by the heuristic mechanism. Verschueren et al. (2003) compared slow and fast responses in causal conditional reasoning tasks with a regression analysis. For the regression analysis probability information (estimated values of $P(q|p)$ and $P(p|q)$) and availability of counterexamples information were used to predict acceptance of inferences. As predicted by their dual-processes approach, fast responses depended only on the probability information while slow responses depended mainly on the counterexamples information. This difference was not due to individual differences between participants, since most participants provided approximately the same amount of fast and slow answers. These results have been confirmed and extended in further experiments. Verbal protocols for example showed that in trials where participants mentioned frequency adverbs (counted as indicator of the heuristic mechanism) as well as a counterexamples (counted as indicator of the analytical mechanism), reasoners mostly started with the frequency adverb and qualified it subsequently with a counterexample, e.g.:

> *Question: "If a predator is hungry, then it will search for prey. A predator is searching for prey. Is the predator hungry?"*
> *Answer: "It* **probably** *is,* **unless** *it hunts to* **feed its** *kids."*

This supports the claim that the heuristic mechanism is primordial and can be overruled. If the indirect pathway in the integrated model would be linked to the heuristic mechanism and the direct pathway to the analytical mechanism, this two process approach could explain why the effects of counterexamples on reasoning were twofold.

The model fit for the sub-model for the noncausal conditionals alone was unsatisfactory. Nonetheless, the direct pathways from availability of exceptional situations to the reasoning variables received significant weights. Furthermore, the paths between availability of exceptional situations, the subjective conditional probability and the overall degree of belief in the conditional also received significant weights – these are the first steps of the indirect pathway. Hence, the first steps in the indirect pathway were stable across conditions (causal and noncausal conditionals). But remarkably, there was no significant effect of exceptional situations on reasoning that was mediated through subjective probabilities – after the direct effect of exceptional situations on the acceptance of MP and MT was partialed out. Due to the rather small sample size (there were twice as many causal as noncausal problems), the results regarding the noncausal conditionals alone should be treated as preliminary. Should this result prove replicable, however, it would contradict the probabilistic account of reasoning as advanced by Oaksford and colleagues (Oaksford et al., 2000; Oaksford & Chater, 2001). This is because Oaksford et al. assume that participants' willingness to accept MP and MT depends on the conditional probability of the consequent, given the antecedent. Whether the conditional itself is of causal or noncausal nature should be irrelevant for the reasoning output.

Varying the plausibility of premises, Rips (2001) declared that humans possess at least two unique ways of evaluating arguments: deductively and in terms of their inductive strength. Contrary to Rips' results, in Experiment 3 only minor differences in acceptance of inferences were observed between two instruction groups that matched Rips' manipulation of deduction versus induction. In the causal path analysis there was even no evidence at all for differences between the instruction groups. Because there were reliable, though small differences between the instruction groups, it seems unlikely that the differences between the instructions were not pronounced enough to activate the different modes of thinking suggested by Rips. Furthermore, the wordings of the questions differed between instruction groups which gave another memory cue to the instructions given at first in the experiment. Then why were the differences between the groups so small respectively absent? Markovits and his colleagues advocate that the ability to *inhibit* background knowledge that is pragmatically relevant to the task but logically irrelevant, is crucial whenever participants are requested to "assume that the premises are true" irrespective of their empirical truth (Markovits & Potvin, 2001; Vadeboncoeur & Markovits, 1999). One support for this general claim constitutes the so-called *paradoxical MP suppression effect* of the generation of an alternative antecedent (Markovits & Potvin, 2001). In several experiments the authors have shown that participants with a standard instruction to assume the truth of the premises tend to deny the MP inference if they have to generate one possible alternative antecedent between the instruction and the reasoning task. They label the effect as paradoxical, because availability of alternative antecedents has been known to suppress AC and DA, but neither MP nor MT. The paradoxical suppression effect was not observed when an alternative antecedent was explicitly presented. Markovits and Potvin (2001) explain this finding by assuming that the active generation of an alternative antecedent reactivates the complete knowledge participants have about the conditional. This reactivation "will reduce the effects of the previous inhibition, and will lead to a greater tendency to deny the MP inference" (Markovits & Potvin, 2001, p. 738). In Experiment 3 there was no intervening task between instruction and reasoning. Yet, there was a cover story positioned between instruction and the conditionals with the reasoning tasks. Maybe, these vivid pseudo-natural stories activated relevant knowledge in semantic memory, which according to Markovits and Potvin (2001) would have reduced the inhibition needed to follow the logic instruction. If this line of thought is correct, it might explain why there were only minor differences between induction and deduction instruction. Then, the experiments would add further support to the conclusion that "this process of selective inhibition is quite fragile" (Markovits & Potvin, 2001, p. 744).

Differences were not only observed between causal and noncausal conditionals. Also, the structure of the description of the causal relationship affected how participants understood and reasoned from the conditionals. The most remarkable difference between forward and backward causal conditionals was observed in the reasoning patterns: There was clear evidence that participants sometimes mentally reversed forward wordings into backward forms. Interestingly, the reversal effect was not restricted to participants that judged the plausibility of the argument, but was observed in the deduction group as well. The reversal constitutes empirical evidence for the often supposed assumption that forward wordings are the natural way to describe a causal relationship. The reversal effect was predicted by theoretical and empirical work by Gavanski and Hui (1992) on probability judgments in category-feature relations. Analogous to category-feature relations, it can be assumed that causes and effects are represented in memory in an asymmetrical way: Causes are recorded with their possible effects, but there is no easy access from effects, given as cues, to their possible causes. Following up on this hypothesis promises interesting insights into knowledge representation in language and in memory, and how the structure of representations can impact reasoning. Nonetheless, in this dissertation these questions will not be pursuited any further.

What are the major results of Experiments 1 to 3? For causal conditionals the experiments yielded simultaneous support for two rival theories: the modified mental model approaches and the probabilistic approach to reasoning. It was proposed that these approaches might be reconciled within a dual-process-theory for causal conditionals that assumes a heuristic mechanisms based on probabilities and an analytical mechanism operating with the availability of counterexamples (Verschueren et al., 2003). Whether a causal conditionals was phrased as "if cause, then effect" or as "if effect, then cause" modified the answers to the reasoning task. This finding constitutes evidence that knowledge representation and usage in memory and language interact with reasoning. Contrary to prior expectations based on results by Rips (2001), the comparison between two instructions, one stressing the logical nature of the task another requesting a plausibility judgment, yielded only a small logical validity effect for AC and MP. It was suspected that this nearly absence of an instruction effect is due to difficulties in the inhibition needed to follow the deduction instruction caused by the vivid pseudo-natural cover stories. Conditionals in which antecedent and consequent were not connected through a causal relation differed from causal conditionals: The data from the noncausal conditionals could not successfully be modeled with the integrated framework. Aside from the path model, the noncausal conditional received clearly lower estimates than the causal ones in both believability measures, subjective conditional probability and the degree of belief. The next experiments will follow up on this finding by examining whether the presence of a causal link between antecedent and consequent still affects the degree of belief and the reasoning process when the conditional probabilities are held constant over causal and noncausal problems.

4 Experiments 4 and 5 (Second Experimental Series)

The semantic link hypothesis suggested previously holds that the presence of a semantic relation between antecedent and consequent events contributes to the conditional's accredited degree of belief. Experiment 2 yielded some support for this view: Conditionals embedded in a causal cover story were estimated to have a higher degree of belief than conditionals embedded in a neutral cover story. The probabilistic approach to the meaning of conditionals could explain this effect by assuming that it is mediated by differences in the assigned subjective conditional probability of the consequent given the antecedent. As a matter of fact, Experiment 1 showed that causal conditionals received higher estimates of the subjective conditional probability than noncausal ones. The previous results are therefore consistent with the semantic link hypothesis but as well with the conditional probability approach, therefore a stricter empirical test seems desirable. Experiment 4 and 5 provide this test: They investigate whether the mere presence of a causal link between antecedent and consequent affects the degree of belief and the reasoning process - independently from the conditional probability. To achieve this goal, the conditional probability of the consequent given the antecedent, *P(q|p)*, was held constant for the comparison of causal and noncausal conditionals. This was possible through inclusion of explicit frequency information about occurrences of the four possible cases (*pq*, *p¬q*, *¬pq*, and *¬p¬q)* into the problems. Introducing explicit frequency information into the problems allows furthermore testing different theories of the meaning of conditionals and the reasoning from conditionals against each other.

Experiment 4 investigated the influence of explicit frequency information on the degree of belief in the conditional. Experiment 5 examined how explicit frequency information affected the reasoning from conditionals. In both experiments causal conditionals were compared to noncausal conditionals.

4.1 Experiment 4: Degree of Belief in the Conditional

Experiment 4 addressed the question how explicit frequency information provided along in the cover stories affected the degree of belief in the conditional for causal and noncausal conditionals.

Edgington (1995) assumes that the degree of belief a listener should have in a conditional statement depends on the conditional probability that the consequent *q* is true given the antecedent *p*. This conditional probability, *P(q|p)*, equals the frequency of *pq*-cases relative to all cases of *p,* which is *pq*-cases plus *p¬q*-cases. Therefore, the degree of belief participants assign to a conditional should depend only on the two sorts of *p*-cases, namely on the *ratio* of *pq*-cases to *p¬q*-cases. To test this prediction, frequency distributions were used that had a ratio of *pq*-cases to *p¬q*-cases that was either high (9:1) or low (1:1). As a second factor the frequency of *pq*-cases was varied, it was either high (900) or low (90). The frequency of *pq*-cases was manipulated, because according to predictions of the mental model theory the frequency of true antecedent/true consequent cases plays a major role for the degree of belief in a conditional, as will become evident later on. Crossing the two factors *ratio* of *pq*-cases to *p¬q*-cases (henceforth shortly "ratio") and frequency of *pq* (henceforth sometimes shortly "frequency") orthogonally yielded the four frequency distributions used which are depicted in Table 7.

Table 7: Frequency Distributions Used in Experiment 4 and 5

	HH	HL	LH	LL
pq	900	900	90	90
p¬q	100	900	10	90
¬pq	500	100	950	910
¬p¬q	500	100	950	910

Note.
HH = high frequency of *pq* and high ratio;
HL = high frequency of *pq* and low ratio;
LH = low frequency of *pq* and high ratio;
LL = low frequency of *pq* and low ratio

The conditional probability account, e.g. by Edgington (1995), in its present form cannot account for *biconditional* readings of conditional statements. Nonetheless, a biconditional interpretation is the most prevalent consistent pattern in conditional reasoning tasks as well as in truth table evaluation tasks (Marcus & Rips, 1979; Rijmen & De Boeck, 2003). Can the probability account be extended to biconditional interpretations? A biconditional reading of a conditional can be phrased as "if *p* then *q*, and if not-*p* then not-*q*". A parsimonious expansion of the conditional probability account to biconditional interpretations is to assume that participants consider two conditional probabilities in this case, i.e., $P(q|p)$ and $P(\neg q/\neg p)$. Because $P(\neg q/\neg p)$ was held constant in the design at 0.5, this account only predicts a main effect of ratio as does the conditional probability account with conditional interpretation. But there is another way to paraphrase a biconditional which can be described as "if *p* then *q*, and if *q* then *p*". With this phrasing, $P(q|p)$ and its reverse probability $P(p|q)$ appear relevant. The reverse conditional probability varied in the design. If participants pay attention to the reverse probability, a main effect of frequency would result as is evident from the values of the reverse conditional probability, $P(p|q)$, depicted in Table 8 for each frequency table condition of the design.

According to the mental model theory a complete representation of a conditional consists of three models (MM_3) or of two mental models in case of a biconditional interpretation (MM_2):

(MM_3)	(MM_2)
$p\ q$	$p\ q$
$\neg p\ \neg q$	$\neg p\ \neg q$
$\neg p\ q$	

According to Johnson-Laird, Legrenzi, Girotto, Legrenzi, and Caverni (1999) who expanded the mental model theory to probabilities, participants estimate the probability of a statement by summing up the probabilities of the individual models that make the statement true. By default each model receives equal weight, but if explicit frequency or probability information is at hand, participants will attach this information to individual models. For example, in a biconditional interpretation (MM_2) these are the models *pq* and $\neg p \neg q$. For a probability or believability estimate, participants should sum up the frequencies of these two cases and divide it by the whole sample (here $N = 2000$). The same line of reasoning applies to a conditional interpretation (MM_3), in which participants represent additionally $\neg pq$ as a third possible model. The believability estimate then equals the sum of the frequencies of *pq*, $\neg p \neg q$ and $\neg pq$ cases divided by the sample size *N*. This is equivalent to 1 minus $P(p\neg q)$. A conditional interpretation according to the mental model theory predicts a *negative* frequency effect with high frequency of *pq* cases associated with a low estimate of believability of the conditional together with an interaction of frequency with ratio. This is because the condition HL should be judged less believable than the rest of the conditions. According to the mental model account a biconditional interpretation (MM_2) yields the prediction that the HH condition will be judged more believable than the rest of the conditions which should not be distinguishable from each other, this should lead to a main effect of frequency and an interaction between frequency and ratio.

Furthermore, the mental model approach assumes that participants will often represent the conditional in a simple way with only one explicit model, leaving the other model(s) implicit.

[p] q
...

This so called "initial model" represents the explicitly only the *pq*-case. The square brackets around *p* indicate that the *p*'s are represented exhaustively: there are no cases of *p* other than those connected with *q*. The dots represent an implicit model which stands for the mental footnote that other mental model are possible. Participants often seem to ignore the implicit model, e.g., due to a lack of working memory capacity (Barrouillet, 1996; Barrouillet, Grosset, & Lecas, 2000; Johnson-Laid & Byrne, 2002; Markovits, Doyon, & Simoneau, 2002). If participants estimate the believability of a conditional only from the explicit model of the initial model (mental model without elaboration, shortly MM_1), then this judgment will depend uniquely on the frequency of *pq*-cases.

Table 8: Exact Numerical and Qualitative Predictions from Different Theories

Theoretical Account	Prediction per Condition				Qualitatively
	HH	HL	LH	LL	
Conditional probability, $P(q/p)$	0.9	0.5	0.9	0.5	ratio effect
Reverse conditional probability, $P(p/q)$	≈ 0.65	0.9	≈ 0.1	≈ 0.1	frequency effect [14]
MM_1 (without elaboration)	0.45	0.45	0.05	0.05	frequency effect
MM_3 (conditional)	0.95	0.55	0.995	0.96	negative frequency effect, frequency *x* ratio
MM_2 (biconditional)	0.7	0.5	0.52	0.5	frequency effect, frequency *x* ratio

Note. MM_1 = mental model without elaboration; MM_3 = mental model with conditional elaboration; MM_2 = mental Model with biconditional elaboration

The paradigm and the frequency distributions are identical to a series of experiments by Oberauer and Wilhelm (2003a) and similar to experiments by Evans et al. (2003). In surprisingly strong agreement, both data sets supported the conditional probability approach and to a lesser extent the mental model approach without elaboration (MM_1). There was no support for elaborated mental models, i.e. MM_3 or MM_2. Experiment 4 will try to expand these results into a new domain (causality) and investigates whether the presence of a causal link between antecedent and consequent contributes to the believability of a conditional statement – independently from the conditional probability.

[14] Consideration of the reverse probability alone would additionally predict an interaction of frequency and ratio, but it is likely that in a biconditional reading participants would consider both probabilities $P(q/p)$ and the reverse $P(p/q)$. Without additional knowledge how participants combine the two probabilities, no further prediction than a main effect of frequency of *pq* seems warranted.

Mandel and Lehman (1998) demonstrated that in contingency judgments participants display a general bias towards weighting tests of sufficiency (related to $P(q/p)$) more heavily than tests of necessity (related to the reverse probability $P(p/q)$). Thus, it could be expected that even with a biconditional interpretation participants who follow the conditional probability approach might weight $P(q/p)$ more heavily than the reverse probability $P(p/q)$.

4.1.1 Participants

The experiment was conducted through the internet from November 2002 to October 2003 in the Web-Lab of the University of Potsdam (http://w-lab.de). The following cases were discarded from all analyses: participants who did not provide answers to all items, participants who answered *Yes* to the question "Did you already participate in this experiment?", and participants using an IP-number that had been used before in the same experiment. After the application of these criteria, 377 participants remained. There were 140 men and 234 women (3 participants did not respond to this question). The mean age was 27.8 years (SD 8.8) with an overall range of 14 to 58.

4.1.2 Materials

Every problem consisted of a cover story, followed by the introduction of a conditional statement, followed by explicit frequency information about the four possible cases pq, $p\neg q$, $\neg pq$, and $\neg p\neg q$. For example:

> *A laboratory in Australia has recently discovered a new allergic disease in dogs. The new disease has been named Midosis. Since that time the researchers have discovered many characteristics of Midosis. Among other things, the scientists have detected that the disease makes an affected dog's blood produce the formerly unknown substance Xathylen.*
>
> *Sara is a practicing veterinarian. She assumes that it generally holds that:*
> *'If a dog suffers from Midosis,*
> *then one finds Xathylen in its blood'*
>
> *For 2000 dogs that have recently been investigated in the laboratory, it is known that:*
> *90 dogs suffered from Midosis and had Xathylen in their blood.*
> *10 dogs suffered from Midosis and had <u>no</u> Xathylen in their blood.*
> *950 dogs did <u>not</u> suffer from Midosis and had Xathylen in their blood.*
> *950 dogs did <u>not</u> suffer from Midosis and had <u>no</u> Xathylen in their blood.*

The cover stories and conditionals were the same as had been used before in Experiments 1 to 3. As a new element explicit frequency information of relevant cases was added to each item. A conditional statement was either embedded in a cover story that introduced the antecedent as a cause for the consequent (causal conditional) or provided no meaningful connection between antecedent and consequent (noncausal conditional). As in Experiments 1 to 3, nine different contents were used. Because no systematic effects of order of terms in the conditional ("if Midosis, then Xathylen" versus "if Xathylen, then Midosis") were found in the previous experiments, order of terms was not varied systematically. Only order b) was used (compare material section of Experiments 1 to 3, pp. 12-14).

Each participant received four problems each with a different frequency distribution. The four different frequency distributions resulted from crossing the two factors frequency of pq-cases and ratio of pq-cases to $p\neg q$-cases (conditions HH, HL, LH, and LL, see Table 7, p. 29). The contents of the four problems were randomly selected and randomly assigned to the frequency distributions. Problems were presented in random order for each participant, each on a separate screen. Causal structure (i.e., presence versus absence of an explicit causal link) was realized between participants: A participant received either only causal conditionals or only noncausal conditionals.

4.1.3 Procedure

Each problem was followed by four questions that were absolutely identical to the questions used in Experiment 2. The first question asked for an estimate of the degree of belief in the conditionals, an estimate of the probability of the conditional being true, that is, $P(p \rightarrow q)$:

> *"What do you think how likely it is that Sara's statement holds true?"*

The second question was a rating of the perceived causal strength. The last two questions assessed the availability of exceptional situations (cases of *p* but *not-q*) and of alternative situations (cases of *q* but *not-p*). For details see "Procedure" of Experiment 2 at page 14, an example of an original item (in German) can be found in Appendix C.

4.1.4 Results

Mean estimates of the degree of belief in the conditional were submitted to an analysis of variance with frequency of *pq* and ratio as within factors and causal structure (causal vs. noncausal) as between participants factor. Main effects of both ratio, $F_{(1, 375)} = 185.0$, $p < 0.001$, $Eta^2 = 0.33$, and frequency of *pq*, $F_{(1, 375)} = 93.5$, $p < 0.001$, $Eta^2 = 0.20$, were significant. The probability of the conditional was estimated to be higher if the ratio was high (54.0) rather than low (34.1). Conditionals were also judged to be more likely to be

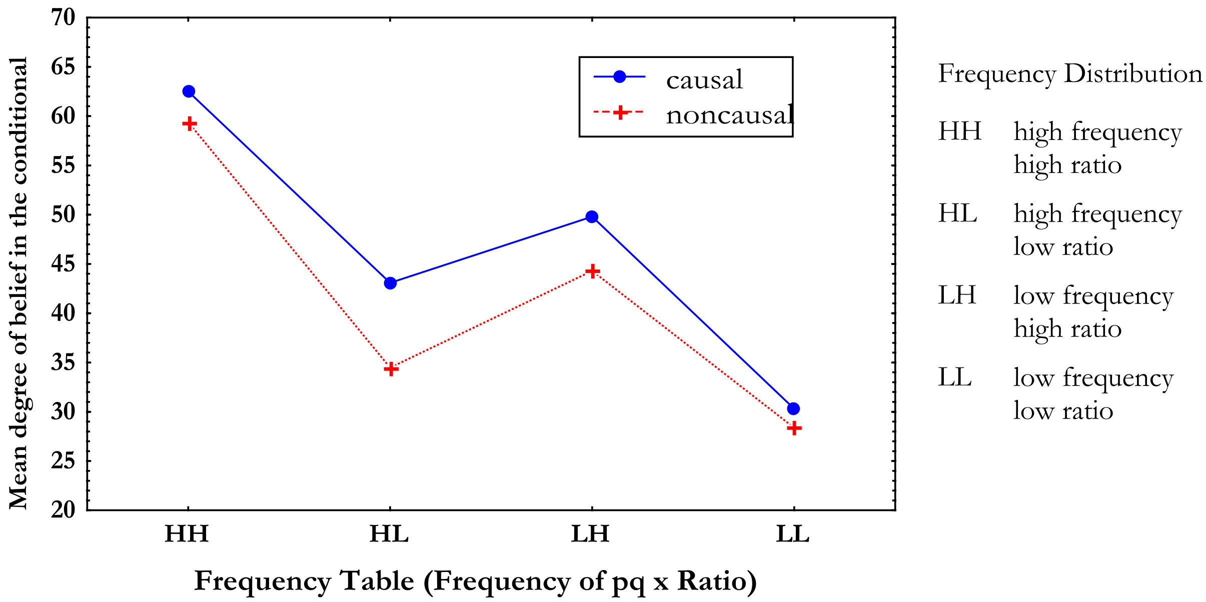


Figure 4: Mean estimated degree of belief in the conditional for each frequency table condition

true if the frequency of *pq* was high (49.9) rather than low (38.2). Mean estimates for each experimental frequency condition, separately for causal and noncausal conditionals, are depicted in Figure 4.

Causal conditionals were estimated to be more believable (46.4) than noncausal conditionals (41.7), yielding a significant main effect of causal structure, $F_{(1, 375)} = 5.1$, $p < 0.05$. A biconditional interpretation according to the mental model theory (MM_2) predicts an interaction between ratio and frequency of *pq*: The ratio effect should be stronger if the frequency of *pq* is high rather than low (compare Table 8, p. 31). This interaction was indeed significant, $F_{(1, 375)} = 4.08$, $p < 0.05$, $Eta^2 = 0.01$. But as indicated by a three-way-interaction between frequency of *pq*, ratio, and causal structure, $F_{(1, 375)} = 4.2$, $p < 0.05$, $Eta^2 = 0.01$, the interaction predicted by MM_2 was found for the noncausal conditionals, but not for the causal problems. This finding could be explained if participants are more likely to interpret noncausal conditionals than causal conditionals as biconditionals. Indeed, in Experiment 3 biconditional reasoning patterns were more frequent in noncausal than in causal conditionals. Another way to interpret the three-way-interaction depicted in Figure 4 is that participants rely on the presence versus absence of a causal link for estimating their degree of belief in the conditional if the two cues from the frequency information (frequency of *pq*, ratio) are in conflict with each other. There is a strong causal link effect in the conditions with high frequency of *pq* and low ratio as well as vice versa (low frequency of *pq* with high ratio) but no causal link effect if either both frequency of *pq* and ratio are high respectively both are low.

Analysis of the means of the estimates of the probability of the conditional provides an incomplete view of the data. The most frequent single value assigned to the degree of belief (the modal response) in each condition equals the ratio: 90 in conditions HH and LH, and 50 in conditions HL and LL. These modal responses account for about a third to a half of the answers. Figure 5 shows a pronounced second peak around the value of 5 for the conditions with low frequency of *pq*, which accounts for 32 and 41 percent of the answers.

The obviously non-normal distributions render an ANOVA a sub-optimal way to analyze the data. Inspection of Table 9 reveals that 64 to 87 % of the answers correspond to one of the stated peaks of the distributions. Loosely speaking, one could refer to the three peaks as expressing a low ("0 to 10"), medium ("45-55"), and high degree of belief ("90-100") in the conditional. It seems warranted to compute a separate analysis for each of the peaks. This multiple testing leads to an increased risk of committing an alpha error which was accounted for by using a stricter criterion for significance: for each of the following tests an alpha criterion of 0.01 was used.

Table 9: Relative Frequencies of Believability Intervals (Low, Medium, High) for Causal and Noncausal Conditionals

Answer	HH	HL	LH	LL
0-10	11.2 / 13.3	18.8 / 31.7	31.5 / 40	40.6 / 43.9
45-55	18.3 / 17.2	58.9 / 53.9	13.2 / 13.3	38.6 / 40
90-100	34.5 / 33.3	6.6 / 1.1	34.5 / 29.4	4.1 / 2.8
Sum	64 / 63.8	84.3 / 86.7	79.2 / 82.7	83.3 / 86.7

Note. First entry in each cell is for causal, the second one for noncausal conditionals.

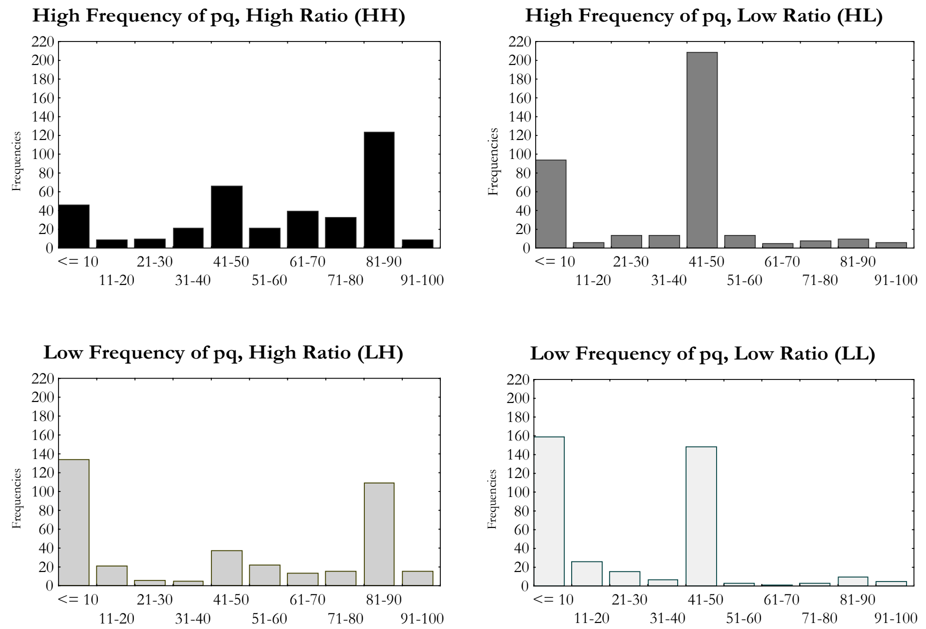


Figure 5: Frequency distributions of intervals of degree of belief in the conditional

The conditional probability approach predicts that participants assign a degree of belief of approximately 90 to the conditions with high ratio (HH and LH) and a belief of approximately 50 to the conditions with low ratio (HL and LH), as shown in Table 8 (p. 31). For this approach, analyses of the medium and high degrees of belief are especially relevant. A log-linear analysis (Wald's test) with frequency of *pq*, ratio, and causal structure as factors was performed with frequency of estimates of high degree of belief in the conditional (i.e., "90-100") as dependent variable. As predicted by the conditional probability account, a high degree of belief was almost never assigned if the ratio was low (mean relative frequencies were 33 % and 4 % for high and low ratio, respectively), yielding a strong main effect of ratio, Wald $X^2(1) = 118.8$, $p < 0.001$. Additionally, there was a trend that participants assigned high degrees of belief more frequently to causal than to noncausal conditionals (mean relative frequencies were 20 % for causal and 17 % for noncausal problems).

The same analysis was performed for the medium level of believability (i.e., "45-55"). As predicted by the conditional probability approach the reverse pattern emerged: A medium level of believability was more frequent in the conditions with low ratio rather than high ratio (48 % and 16 % for low and high ratio, respectively), indicated by a main effect of ratio, Wald $X^2(1)= 165.2$, $p < 0.001$. Additionally, medium levels of belief were assigned more frequently to conditionals with high frequency of *pq* than with low frequency (37 % and 26 % for high and low frequency, respectively) as predicted by the mental model approach without elaboration (MM_1). For the two analyses on the high and medium degrees of belief no other main effect or interaction proved significant. For the exact relative frequencies for the answers in each condition for causal and noncausal conditionals please compare Table 9.

As already mentioned, in the conditions with low frequency of *pq* there was a second peak around 5 which was predicted by the mental model approach without elaboration (MM_1). MM_1 predicts an estimate of approximately 5 to conditions with low

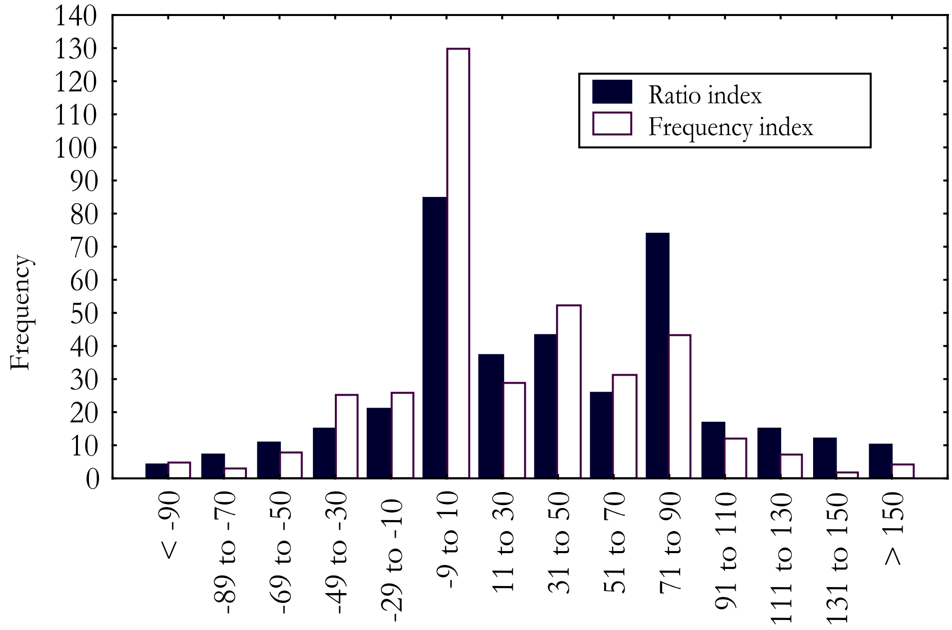


Figure 6: Distributions of ratio and frequency index

frequency of *pq* (conditions LH and LL) and of approximately 45 if the frequency of *pq* is high (HH and HL). The frequencies of estimates expressing a low degree of belief ("0 to 10") were submitted to a log-linear analysis identical to the previous with frequency of *pq*, ratio, and causal structure as factors. The analysis yielded a main effect of frequency of *pq*, as predicted by MM_1, Wald $X^2(1) = 76.8, p < 0.01$: Low degrees of belief were assigned in about 19 % of the cases if the frequency was high, but in 39 % of the problems if the frequency was low. Furthermore, the analysis indicated that a low degree of belief was more often assigned to noncausal conditionals (32 %) than to causal conditionals (26 %), Wald $X^2(1) = 71.9, p < 0.01$. Furthermore and unpredicted by MM_1 there was a main effect of ratio as well, Wald $X^2(1) = 20.7, p < 0.01$: A low degree of belief was more often assigned to conditionals with a low ratio than with high ratio (34 % versus 24 %).

To investigate individual interpretations of the conditional in Experiment 4, two indices were computed that reflect the effect of ratio and of the frequency of *pq* respectively on the estimate of the degree of belief given by a participant. The same indices have been used by Oberauer and Wilhelm (2003a). The ratio index was defined as the probability estimate of (HH+LH) – (HL+LL). The frequency index was defined as (HH+HL) – (LH+LL). High ratio indices together with frequency indices around zero should occur in participants that interpret the conditional as a conditional probability. The reverse should be the case for participants who interpret the conditional according to MM_1: high frequency indices combined with ratio indices around zero. The other combinations should be observed only rarely. Indeed, in Oberauer and Wilhelm's data the two indices were negatively correlated ($r = -0.37$). Here, the two indices also correlated negatively with each other for noncausal conditionals, $r = -.25$ ($p < 0.01$, $n = 180$) as well as for the causal conditionals ($r = -0.19, p < 0.01$, $n = 197$). The distributions of the two indices are depicted in Figure 6.

The ratio index shows a clear bimodal distribution with a peak around 0 and another one around 80. These two peaks are consistent with the interpretation that most participants can be assigned into one of two groups: One group of participants who do not take the ratio into consideration for estimating the degree of belief in the conditional, these participants have a ratio index close to zero. The second group is strongly influenced by the ratio and has a ratio index around 80. A ratio index of 80 arises when the degree of belief estimates perfectly match the ratio of *pq* to *p¬q* times 100, (90+90) – (50+50) = 80.

The distribution of the frequency index does not show a bimodal pattern. The majority of participants has a value around 0 that indicates that these participants do not take the frequency of *pq* into account when estimating the degree of belief. A frequency index of 80 would arise when a participant's estimates match the percentage of *pq*-cases: (45+45) – (4.5+4.5), but only a small fraction of participants have a frequency-ratio close to 80. Table 10 shows a classification of participants, in which two groups are of theoretical interest: those that take into consideration the ratio of *pq* to *p¬q* while ignoring the frequency of *pq*, and those who make their estimates depend on the frequency of *pq*, but ignore the ratio. In Table 10 the cutoff criterion for both indices was set arbitrarily at 40. Overall 138 of 377 participants (35 % and 39 % for causal and noncausal conditionals, respectively) give consistent estimates that are influenced by the ratio, but not of the frequency of *pq*. Answers from 84 participants (21 % and 23 % for causal and noncausal conditionals, respectively) show a consistent pattern influenced by the frequency of *pq*, but not by the ratio.

Causal conditionals received generally higher estimates of degree of belief than noncausal conditionals. Within every frequency condition the conditional probability, *P(q|p),* was held constant, therefore this difference in believability between causal and noncausal conditionals is strong support for the hypothesis that the presence of a semantic link between antecedent and consequent in a conditional contributes to its believability (semantic link hypothesis). Can this semantic link effect be explained by how much *causal strength* participants attribute to the relationship described in the conditional statement? Question 2 asked participants whether they think that the statement describes a causal relationship (0 = No) and if they did so they estimated the strength of this causal relationship on a 5-point-rating scale. Ratings of causal strength were submitted to an analysis of variance with frequency of *pq* and ratio as within participants factors, and causal structure as between participants factor. Although there were clear main effects of frequency of *pq,* $F_{(1,385)} = 18.7$, $p < 0.01$, $Eta^2 = 0.046$ (means were 2.3 vs. 2.0 for high and low frequency of *pq*, respectively) and ratio, $F_{(1,385)} = 49.4$, $p < 0.01$, $Eta^2 = 0.11$ (means were 2.5 vs. 1.8 for high and low ratio, respectively), the main effect for causal structure was only marginally significant ($F_{(1, 385)} = 3.9$, $p = 0.05$, $Eta^2 = 0.01$, means were 2.3 vs. 2.0 for causal and noncausal conditionals, respectively). In Experiment 2 the causal structure had a large effect on the perceived causal strength: Ratings for

Table 10: Classification of Participants According to the Ratio and Frequency Indices

	Ratio index ≤ 40	Ratio index > 40	Sum
Frequency index ≤ 40	66 / 54	68 / 70	134 / 124
Frequency index > 40	42 / 42	21 / 14	63 / 56
Sum	108 / 96	89 / 84	197 / 180

Note. Cell entries are numbers of participants. First entry is for causal conditionals (*n* = 197), second for noncausal conditionals (*n* = 180). Overall *N* = 377.

causal conditionals were actually twice as high as for the noncausal conditionals (means were 3.0 vs. 1.5 for causal and noncausal conditionals, respectively). Explicit frequency information as given in Experiment 4 seems to reduce but not to override completely the effect of the presence versus absence of a causal link in the cover story on the perceived causal strength and might be partially responsible for the difference in the degrees of belief assigned to causal and noncausal conditionals.

Causal Strength estimates in Experiment 2 have been made without any further knowledge other than the cover story and the conditional statement. These ratings can therefore be used as an estimate of the *causal power* participants *spontaneously attribute* to the causal mechanism in the problem. Can this causal power of a mechanism predict the degree of belief participants attribute to a problem in Experiment 4 where explicit frequency information was added to the cover stories? Ratings of the estimated causal strength from Experiment 2 and assigned degrees of belief in Experiment 4 correlate over problems ($r = 0.28$, $p < 0.05$, $N = 72$). But this correlation holds only for the noncausal problems ($r = 0.35$, $p < 0.05$, $n = 36$), not for the causal problems ($p > 0.5$).

4.1.5 Discussion

The ratio of *pq*-cases to *p¬q*-cases had a large impact on the degree of belief participants assigned to a conditional, this supports the conditional probability approach. To a lesser extent the frequency of *pq*-cases also affected the estimated degree of belief, this can count as evidence for the mental model approach without elaboration (MM_1). The results yielded no support for the mental model approach with a fully fleshed-out conditional interpretation (MM_3). The mental model approach with a biconditional interpretation (MM_2) was supported only for the noncausal conditionals. These findings clearly replicate the effects of frequency manipulations on the degree of belief reported by Oberauer and Wilhelm (2003a) and Evans et al. (2003) - with the exception of the support for MM_2 in the noncausal conditionals. Two groups of participants were identified: one group who consistently take into account the ratio but ignore the frequency of *pq* when estimating the degree of belief and one group with the reverse pattern: These participants base their answers on the frequency of *pq* while ignoring the ratio. A possible interpretation is that the first group (approx. 40 % of participants) follows consistently the conditional probability account and the second group (approx. 20 %) the mental model approach without elaboration. The classification of participants explains about 62 % of the answers for noncausal conditionals, but only 56 % of answers for causal conditionals. Although this difference does not reach significance in a *Chi²*-test, X^2 (1) = 1.58, $p_{\text{one-sided}} = 0.12$, it might be interpreted as a possible hint that for the causal conditionals other aspects aside from the frequency information (e.g., believability or familiarity[15] of the described causal relation), are influential with respect to the believability of the conditional statement. Thus, the provided explicit frequency information might have been more effective for noncausal conditionals. A simple alternative explanation is the idea that maybe participants behaved more consistently in the noncausal problems.

[15] Although all described causal relations were fictitious they were similar to causal relations that participants are familiar with.

Conditionals embedded in a causal cover story received higher degrees of belief than conditionals embedded in a neutral cover story. This finding constitutes strong support for the semantic link hypothesis. Because contrary to Experiment 2 the difference in degree of belief between causal and noncausal conditionals cannot be mediated through the conditional probability of the consequent given the antecedent, since this conditional probability was held constant within each frequency condition. But the support for the semantic link hypothesis is not unambiguous: Causal conditionals were only judged to be more believable than noncausal conditionals if ratio and frequency of *pq* were in conflict with each other, e.g., because the ratio was high but the frequency of *pq* was low or vice versa. There was no significant effect of the causal link if ratio and frequency of *pq* agreed, that is both were either high or both were low. It should be noted though that this pattern of results can also be due to more biconditional interpretations according to MM_2 with the noncausal conditionals. The latter explanation is supported through the finding that in Experiment 3 more participants showed biconditional reasoning patterns in noncausal than in causal conditionals. Furthermore, according to the conditional probability approach participants base their believability estimates on the subjective conditional probability of the consequent given the antecedent. Objective conditional probabilities were controlled for experimentally, but this does not necessarily mean that the subjective conditional probabilities were effectively controlled too.

4.2 Experiment 5: Reasoning

Experiment 5 investigates how participants reason from conditional problems with explicitly provided frequency information. The same four frequency tables as in Experiment 4 were used, crossing the factors frequency of *pq* and ratio. For the four frequency tables please see Table 7, p. 29.

4.2.1 Predictions

From a logical viewpoint, the acceptance of a conclusion in a deductive task depends only on the premises given. Neither the cover stories nor the frequency information should affect the reasoning performance. From a psychological viewpoint it is highly implausible that participants won't be affected by these informations provided along with the problem. One argument comes from pragmatics: It would be odd and even misleading to provide information which is not relevant for the task at hand. In the *modified mental model approaches* introduced and discussed above (Markovits & Barrouillet, 2002; Barrouillet & Lecas, 1998; Schroyens, Schaeken, & d'Ydewalle, 2001; Schroyens & Schaeken, 2003), explicit frequency information can be considered as part of the (acquired) background informations that are automatically activated and used in reasoning. The background knowledge can be actively inhibited, e.g., under deductive instructions, but only to a certain extent because the inhibition is very resource-demanding. According to these approaches, major and minor premise serve as cues that automatically activate all knowledge associated with them. The putative conclusion will be accepted if no counterexample is retrieved from memory or if the retrieved counterexample is not very likely to be true (Schroyens et al., 2001). A counterexample is a case in which the minor premise is true, but the provisional inference is definitely false, i.e. *p¬q*-cases for MP and MT and *¬pq*-cases for AC and DA. Furthermore, according to Schroyens et al., 2001, and contrary to the traditional mental model theory (Johnson-Laird, 1983) the search for a counterexample is performed if and only if a participant is uncertain about the truth of the provisional conclusion. Johnson-Laird et al. (1999) have extended the mental model theory to probabilities: By default, all mental models receive equal weights. But if explicit frequency information is at hand, mental models can receive different weights. In that situation a model receives a weight that equals the frequency associated with the model divided by the sum of the frequencies of all possible cases. According to the modified mental model approach whenever participants are uncertain of the truth of the conclusion, the frequency of empirical counterexamples should directly and negatively relate to the acceptance of the inference. The number of counterexamples varied in the experimental conditions, although not in a systematic manner. The first row of Table 11 lists the number of counterexamples to MP and MT in each frequency distribution condition, and the second row contains the number of counterexamples to AC and DA.

Table 11: Number of Counterexamples in the Sample of 2000 Cases

Inferences and counterexample	HH	HL	LH	LL	Qualitative prediction on acceptance
MP + MT (*p¬q*)	100	900	10	90	HL < rest
AC + DA (*¬pq*)	500	100	950	910	main effect of frequency of *pq*

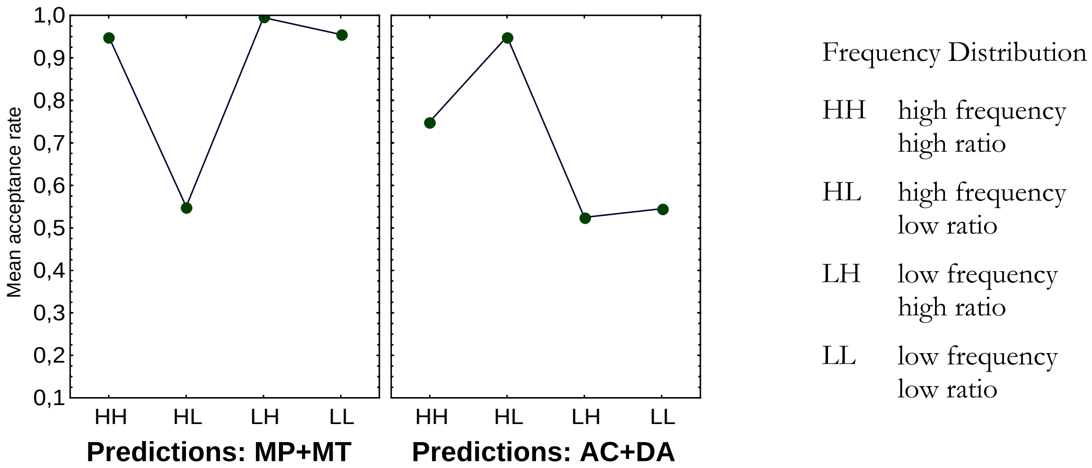


Figure 7: Predictions for acceptance rates according to the modified mental model theories

The modified mental model approach predicts that inferences should be accepted in direct negative proportion to the number of counterexamples to it. Figure 7 illustrates the resulting predictions on the mean acceptance rates of inferences separately for MP and MT as well as AC and DA. Qualitatively, the modified mental model approaches predict for MP and MT that the lowest acceptance rates will be observed in condition HL. For AC and DA this approach predicts a main effect of frequency of *pq*, because the conditions HH and HL with high frequency of *pq* have more empirical counterexamples to AC and DA (i.e., cases of *¬pq*) than the conditions LH and LL. In the model variant of Schroyens and Schaeken (Schroyens et al., 2001; Schroyens & Schaeken, 2003), the two negated inferences MT and DA depend furthermore on the likelihood of considering the false-antecedent/false-consequent contingency (*¬p¬q*). If participants fail to build this model, they will reject MT and DA. Therefore this approach additionally predicts that in all conditions the acceptance rate of MP is higher than of MT and the acceptance rate of AC is higher than that of DA.

In the *probabilistic theory of conditional reasoning* by Oaksford et al. (2000) an inference is accepted in direct proportion to the conditional probability of the provisional conclusion given the categorical premise, e.g., for MP = *P(q|p)* and for MT = *P(¬p|¬q)*. This account yields exact predictions for each inference in every frequency condition, they are given in Table 12 and depicted in Figure 8. But as Oaksford and Chater (2003a, 2003b) have stressed their theory is settled at a computational level describing thus optimal normative behavior. To specify actual human performance, e.g. in a reasoning task, requires further processing assumptions that have not been worked out yet. To account for this, only the qualitative predictions that can be derived from the exact predictions will be considered. The probabilistic

Table 12: Predictions for Acceptance of Inferences from Oasksford et al. (2000)

	Exact				Qualitative
	HH	HL	LH	LL	
MP	0.90	0.50	0.90	0.50	ME Ratio
MT	0.83	0.10	0.99	0.91	HL < rest
AC	0.64	0.90	0.09	0.09	ME Frequency of *pq*
DA	0.50	0.50	0.50	0.50	no effect

Note. ME = main effect; frequency = frequency of *pq*-cases; ratio = ratio of *pq*-cases to *p¬q*-cases

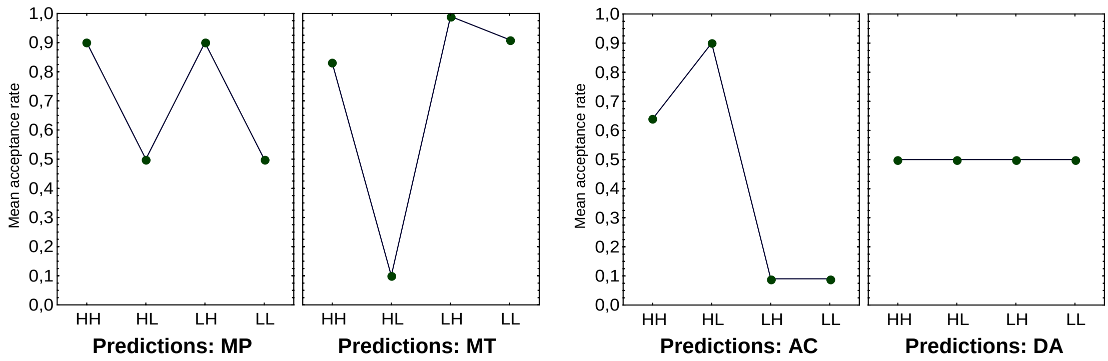


Figure 8: Predictions for acceptance rates according to Oaksford et al. (2000)

account predicts that acceptance rates for MP should depend only on the ratio of *pq*-cases to *p¬q*-cases. Acceptance rates for the MT inference should be especially low in the frequency condition with high frequency of *pq* and low ratio (HL). For AC a main effect of frequency of *pq* is expected. For DA there should be no effect of the frequency manipulation whatsoever.

The modified mental model approaches and the probabilistic theory of Oaksford et al. (2000) yield qualitatively similar predictions for AC and MT, but disagree on the effects that the explicit frequency information will have on MP and DA. Therefore, MP and DA inferences have a special diagnostic role in deciding between these theoretical approaches.

Participants who follow a normative approach and base their answer on the logical status either according to the material implication (conditional) or the material equivalence (biconditional) will show no effect of the frequency information as depicted in Figure 9 for the conditional interpretation. Those participants would accept MP and MT, while rejecting AC and DA (conditional), or accept all four inferences (biconditional) irrespective of the frequency information. Participants answering according to the conditional as material implication would therefore increase overall acceptance rates of

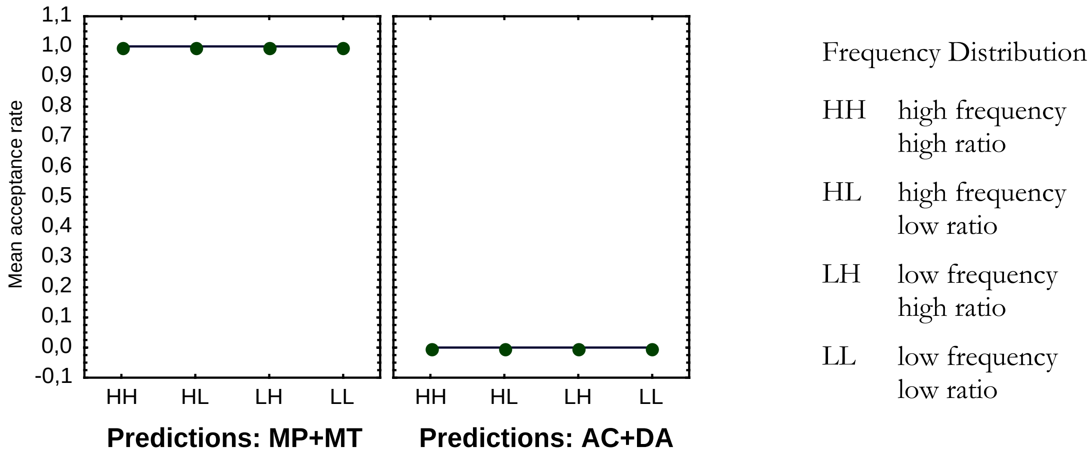


Figure 9: Predictions for acceptance rates following a conditional interpretation

MP and MT, while decreasing acceptance rates of AC and DA. They would not add systematic variance depending on the frequency condition.

The integrated model suggested in section 2 (pp. 3-8) assumes that MP and MT depend on the believability of the conditional which corresponds to the conditional probability of the consequent given the antecedent, that is *P(q|p)*. Therefore, for MP and MT there should be a main effect of ratio on acceptance rates, because in the conditions with high ratio the conditional probability is high and with low ratio it is low. The integrated model can theoretically be extended to account for acceptance of AC and DA inferences: According to the model these inferences are accepted if participants believe that the reverse conditional "if *q* then *p*" is true. The believability of this conditional is equal to the conditional probability of the original antecedent given the original consequent, *P(p|q)*. This reverse conditional probability can be read off from Table 8, p. 31. The reverse conditional probability is rather high in the conditions with high frequency of *pq*, therefore the integrated model predicts a main effect of frequency of *pq* on the acceptance of AC and DA. Table 13 provides an overview over the predictions made by the modified mental model theories, the probabilistic theory of Oaksford et al. (2000), and the integrated model.

Table 13: Overview of Predictions for Acceptance of Inferences

	Modified mental model theories	Oaksford et al. (2000)	Integrated model
MP MT	} HL < rest	ME ratio HL < rest	} ME ratio
AC DA	} ME frequency	ME frequency no effect	} ME frequency

Note. ME = main effect; frequency = frequency of *pq*-cases; ratio = ratio of *pq*-cases to *p¬q*-cases

4.2.2 *Participants*

Experiment 5 was conducted through the internet between September and December 2002. 394 people participated overall, 206 in the group with causal problems and 188 in the group with the noncausal problems. In the causal group there were 120 male and 85 female participants, in the noncausal group there were 109 men and 78 women (the remainder did not respond to this question). Mean age was 26.5 years (SD = 9.5) with an overall range from 14 to 62 years.

4.2.3 *Materials and Procedure*

Experiment 5 used the same problems as Experiment 4, this time followed by four inference tasks: MP-AC-MT-DA or DA-MT-AC-MP. Order of inferences was held constant within a participant. Participants were supposed to judge whether an inference followed with logical necessity from the conditional statement and the minor premise. Instruction and presentation of the inferences were identical to the instruction and the format for the deduction group in Experiment 3 (compare 3.2.3, p. 14). Since in Experiment 3 confidence ratings did not yield any interesting results, they were not assessed in Experiment 5. Every participant received four problems, each with a different frequency distribution. As in Experiment 4 contents were assigned randomly to the frequency distributions, problems were presented in random order, each on a different screen. A participant obtained either only causal or only noncausal problems. An example of an original item (in German) can be found in Appendix C.

4.2.4 *Results*

4.2.4.1 *Acceptance Data*

Acceptance data were submitted to a log-linear analysis for frequencies (Wald's test) with inference form (MP vs. MT vs. AC vs. DA), frequency of *pq* (high vs. low) and ratio (high vs. low), and causal structure (causal vs. noncausal) as factors. Overall acceptances rates were rather low varying between 27 and 44 percent (corresponding values in Experiment 3 were 36 to 85 percent). This shows that participants were largely affected by the explicit frequency information, although the instruction stressed that participants were supposed to judge only whether the conclusion followed from the major and minor premise with logical necessity. Unsurprisingly, acceptance rates varied considerably between inferences (mean acceptance rates were 0.44, 0.36, 0.27 and 0.28 for MP, MT, AC, and DA, respectively), producing a main effect of inference form, Wald $X^2(3) = 134.6$, $p < 0.01$. Inferences from a problem in a condition with a high ratio of *pq*-cases to $p\neg q$-cases were accepted more frequently than from problems with low ratio as shown by a large main effect of ratio, Wald $X^2(1) = 45.0$, $p < 0.01$ (means were 0.38 and 0.30 for high and low ratio, respectively). There was a small but significant main effect for frequency of *pq* as well: fewer inferences were accepted in problems with low frequency than in problems with high frequency of *pq*, Wald $X^2(1) = 6.8$, $p < 0.05$ (means were 0.35 and 0.32 for high and low frequency of *pq*, respectively). Participants were more willing to accept inferences if a causal link was explicitly mentioned (means were 0.36 and 0.32 for causal and noncausal problems, respectively), yielding a main effect of causal structure, Wald $X^2(1) = 14.2$, $p < 0.01$. There were furthermore several significant two-way-interactions: ratio interacted with causal structure, Wald $X^2(1) = 10.9$, $p < 0.01$, and inference form interacted with frequency of *pq*, Wald $X^2(3) = 46.0$, $p < 0.01$ as wells as with ratio, Wald $X^2(3) = 30.4$, $p < 0.01$. These interactions were followed up by separate analyses for each inference.

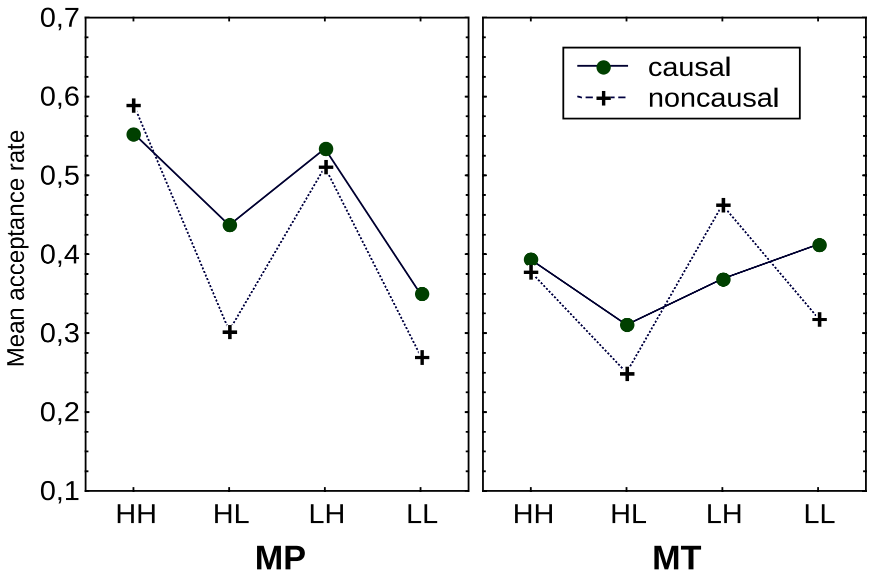
MP inferences were accepted more often if the ratio was high rather than low as indicated by a large main effect of ratio, Wald $X^2(1) = 67.9$, $p < 0.01$ (mean acceptance rates were 0.55 and 0.34 for high and low ratio, respectively, see also Figure 10). Furthermore, more MP inferences were accepted in conditions with high frequency of *pq* (HH and HL), yielding a small main effect of frequency, Wald $X^2(1) = 7.9$, $p < 0.05$ (means were 0.47 and 0.42). Participants were more willing to accept a MP inference from a causal than from a noncausal problem, Wald $X^2(1) = 4.5$, $p < 0.05$ (means were 0.47 and 0.42). The ratio effect was clearly stronger for the noncausal problems (means were 0.55 and 0.29 for low and high ratio, respectively) than for the causal problems (means were 0.54 and 0.39 for low and high ratio, respectively), resulting in a significant interaction between ratio and causal structure, Wald $X^2(1) = 5.7$, $p < 0.05$.

For MT there was a significant main effect of ratio too, Wald $X^2(1) = 10.6$, $p < 0.01$ (means were 0.40 and 0.32 for high and low ratio, respectively). There was a main effect of frequency of *pq*, Wald $X^2(1) = 5.9$, $p < 0.05$, but it runs in the opposite direction as the frequency effect for MP: more MT inferences were accepted in problems with a low frequency of *pq* (0.39) than with high frequency (0.33). This negative frequency effect is probably due to the very low acceptance rates of MT in the HL condition, this finding will be discussed later on. As with the MP inference, there was an interaction between ratio and causal structure, Wald $X^2(1) = 5.6$, $p < 0.05$: The ratio effect was large and reliable for the noncausal problems (means were 0.42 and 0.28 for high and low ratio, respectively), but absent for the causal problems (means were 0.36 and 0.38 for high and low ratio, respectively).

As predicted by the probabilistic theory and the modified mental model theory, the HL-condition with high frequency of *pq* coupled with low ratio showed the lowest acceptance rates for MT, as indicated by a planned contrast comparing condition HL and the other conditions, Wald $X^2(1) = 15.1$, $p < 0.05$. This contrast was significant separately for the causal as well as for the noncausal problems, Wald $X^2(1) = 4.3$, $p < 0.05$ and Wald $X^2(1) = 11.3$, $p < 0.01$. Although the effect seems to be stronger for the noncausal conditionals (means were 0.25 for HL and 0.39 for the other three conditions) than for the causal problems (means were 0.31 for HL and 0.39 for the other three conditions), the corresponding interaction did not reach significance, Wald $X^2(1) < 1.5$.

AC inferences were accepted more often in conditions with high frequency of *pq* than with low frequency (means were 0.34 and 0.20, compare as well Figure 11), as indicated by a large main effect of frequency of *pq*, Wald $X^2(1) = 39.1$, $p < 0.01$. There was a small trend that participants drew more AC inferences if the problem was causal than noncausal (means were 0.29 and 0.25 for causal and noncausal problems), Wald $X^2(1) = 3.8$, $p = 0.05$. No other effect was found to be significant for AC.

Acceptance rates of the DA inference were not affected by the frequency manipulation. The analysis yielded only a significant effect for the causal structure: Participants were more willing to accept DA inferences if they stemmed from causal than from noncausal problems (means were 0.30 and 0.25), Wald X2(1) = 6.4, p < 0.05. Figure 11 displays the mean acceptance rates for AC and DA.



Frequency Distribution

HH	high frequency high ratio
HL	high frequency low ratio
LH	low frequency high ratio
LL	low frequency low ratio

Figure 10: Mean acceptance rates for MP and MT in problems with different frequency distributions

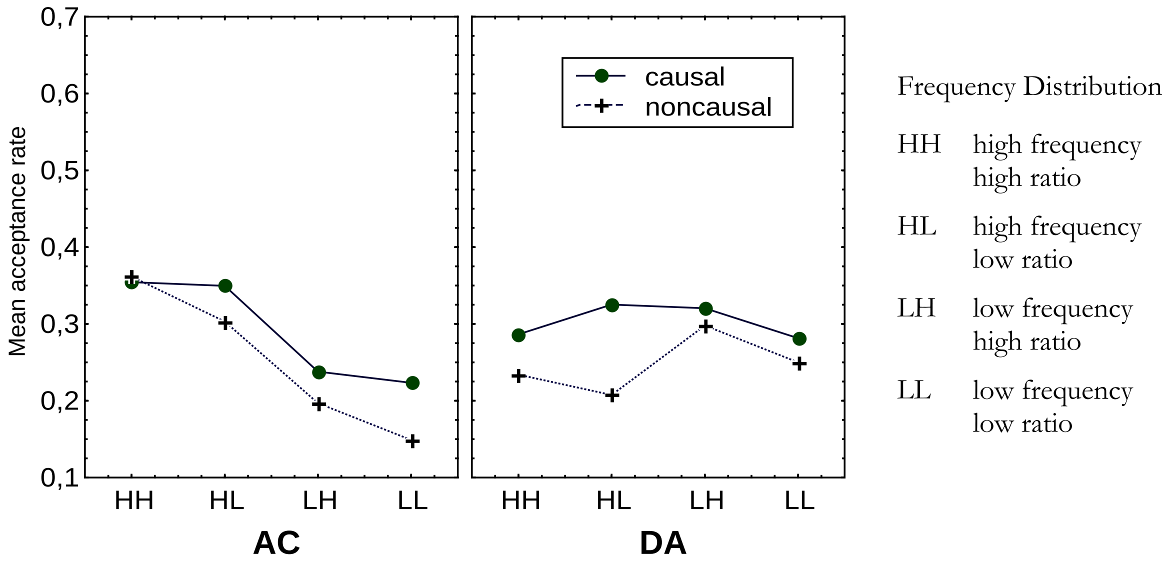


Figure 11: Mean acceptance rates for AC and DA in problems with different frequency distributions

In most kinds of deductive reasoning tasks a logical validity effect is observed: Participants are more likely to accept an inferences if it is logically valid than logically invalid (Schroyens et al., 2003). Is there a logical validity effect in the data from Experiment 5 or does the frequency information containing empirical counterexamples to each inference override the logical validity effect usually found? To approach this question, a regression was computed. Logical status (0 for AC and DA; 1 for MP and MT) and the number of empirical counterexamples in the frequency information were used as predictors for the mean acceptance rate of an inference. Problems were aggregated over semantic contents, leaving 32 cases with the structure inference form (MP vs. MT vs. AC vs. DA) *x* frequency table (HH vs. HL vs. LH vs. LL) *x* causal structure (causal vs. noncausal). The regression yielded significant beta-weights for both predictors: logical status (b = -0.46, t = -3.12, $p < 0.05$) and number of empirical counterexamples (b = -0.38, t = -2.58, $p < 0.05$). The regression explained 50 percent of the variance in the acceptance rates. The results of the regression analysis show that there is still a logical validity effect – even with explicit frequency information of counterexamples. This justifies the conclusion that participants were sensitive to empirical counterexamples as well as logical status of an inference in their decision to accept or reject an inference.

4.2.4.2 Reasoning Patterns

Because every problem was followed by all four inferences, it is possible to investigate the *pattern of answers* to the inferences of each problem solved as has been done for Experiment 2 (3.3.4 Reasoning Patterns, pp. 23-25). In Experiment 5 participants received 1576 problems overall (824 causal and 752 noncausal problems). Table 14 displays a classification of relevant reasoning patterns. Surprisingly, the most frequent pattern observed in the data was the rejection of all four inferences. In 40 % of the problems the participants decided to accept none of the inferences. This is a very surprising finding, especially since this pattern was never observed in Experiment 3. Further important patterns observed were the biconditional and conditional pattern. A pattern of answers was classified as *biconditional* if all inferences were accepted. It was classified as *conditional* if MP and MT were accepted, but AC and DA were rejected. Biconditional and conditional patterns account only for nearly 20 % of the problems solved (means were 9.3 % and 8.9 % for biconditional and conditional pattern, respectively) which is a low number compared to Experiment 3 where these two logically valid systematic answers constituted about 60 % of answers (approx. 40 % biconditional and 20 % conditional). The majority of systematic answers in Experiment 5 belonged into the rejection category.

The frequencies for each pattern were submitted to a log-linear analysis (Wald's test) with causal structure (causal vs. noncausal) and frequency of *pq* and ratio as factors. The *rejection* of all inferences was more frequent if the ratio was low rather than high (Wald $X^2(1) = 16.5, p < 0.01$, means were 45 and 35 % for low and high ratio, respectively). The biconditional pattern was more prevalent in causal than noncausal conditionals, as indicated by a significant main effect of causal structure, Wald $X^2(1) = 6.6, p < 0.05$ (means were 11 % and 7 % for causal and noncausal problems). There was furthermore a main effect of frequency of *pq*, Wald $X^2(1) = 5.9, p < 0.05$: There were more biconditional answers if the frequency was high rather than low (means were 11 % and 7 % for high and low frequency of *pq*, respectively). Conditional patterns were more common with high ratio (12 %) than with low ratio (6 %), Wald $X^2(1) = 15.0$, $p < 0.01$. Additionally there was a small trend that conditional patterns occurred more frequently in the causal (10 %) than in the noncausal problems (8 %), Wald $X^2(1) = 3.3$, $p = 0.07$.

Table 14: Percentages of Reasoning Patterns

	Causal	Noncausal
Reject all	38.23	41.36
Biconditional	11.04	7.31
Conditional	9.95	7.85
One mental model	3.76	4.12
Reverse conditional	1.21	2.13
Other	35.80	37.23

Note. Reject all = MP+MT+AC+DA rejected. Biconditional = MP+MT+AC+DA accepted. Conditional = MP+MT accepted, AC+DA rejected. One Mental Model = MP+AC accepted, MT+DA rejected. Reverse Conditional = MP+MT rejected, AC+DA accepted.

Considering the *believability* of a conditional as a mediator between the frequency manipulation and the answering patterns, provides a possible explanation for the observed effects. If the effects of the frequency information on the reasoning patterns are due to a spill-over from the effects of the frequency information on the believability of the conditional, it would be expected from the probabilistic approach to the meaning of conditionals that participants with a conditional interpretation are sensitive to the ratio and participants with a biconditional pattern to the frequency of *pq* (compare theoretical part of Experiment 4, pp. 29-32). This matches the observed small but significant effects of the frequency information on the reasoning patterns.

Another question raised by the analyses of the reasoning patterns is whether some participants show a consistent pattern across the four problems solved. Only a very small fraction of participants show a persistent pattern that is either conditional (1.5 %) or biconditional (1.3 %) for all four problems. But about a fourth of the participants (25.4 %) rejects all inferences in every problem. Additionally, 33.2 % of participants reject all four inferences in one, or two, or three of the problems solved.

4.2.5 *Discussion*

The participants' willingness to accept inferences depended strongly on the explicit frequency information provided in the problem. The theory that was able to explain most of the major results is the probabilistic approach to reasoning by Oaksford et al. (2000). It correctly predicted a main effect of ratio for MP, low acceptance rates for MT in the condition with high frequency and low ratio (HL), a main effect of frequency of *pq* for AC, and the absence of any effect of the frequency manipulation on the DA inference. The only main finding that the approach fails to explain is the observed small effect of the frequency of *pq* on MP.

The integrated model predicts that MP and MT depend on the believability of the conditional which corresponds to $P(q|p)$, and that AC and DA hinge on the reverse probability $P(p|q)$. This model captures some basic results, e.g., the ratio effect on MP and MT and effect of the frequency of *pq* on AC. But it cannot explain why there is no effect of the frequency of *pq* on DA and why there are small frequency of *pq* effects in the data for MP and MT.

The modified mental model approaches correctly predicted a main effect of frequency of *pq* on AC and a low acceptance of MT in the condition with high frequency and low ratio (HL). But the approaches can't explain the absence of an effect of the frequency of *pq* on DA. Schroyens et al. (2001) furthermore predicted that MP should be accepted more frequently than MT and AC more often than DA, because the latter inferences depend on the successful construction of the false antecedent/false consequent-model ($\neg p \neg q$). MP was indeed accepted more frequently than MT, but AC wasn't accepted more often than DA. It is difficult to evaluate the lack of a difference in acceptance rates between AC and DA, because the finding might be due to a floor effect.

For the predictions of the mental model approaches it is crucial how participants assign weights to mental models. Johnson-Laird et al. (1999) have suggested that the weight given to a mental model equals the frequency associated with the model divided by the sum of the frequencies of all cases. This method yields what may be called "conjunctive weights". However, this is not the only possible solution to the assignments of weights. Oberauer and Wilhelm (2003b) have recently suggested introducing the concept of a *reference frame* into the ontology of mental models. "A reference frame defines explicitly a region in a mental space of possibilities to which a mental model should be interpreted" (Oberauer & Wilhelm, 2003b, p. 4). It seems reasonable to assume that in conditional reasoning tasks the minor premise defines the current reference frame. The weight assigned to the model of counterexamples would equal the frequency of counterexamples divided by the frequency of cases in the reference frame (i.e., all cases matching the minor premise). These weights could be called "relative weights". If the mean expected acceptance rates of inferences are computed from the relative weights instead from the conjunctive weights, predictions emerge that are identical to the predictions of the probabilistic theory of reasoning by Oaksford et al. (2000). A modified mental model approach with relative weights according to reference frames would thus be able to explain all major results found in Experiment 5. But there is no experimental evidence of reference frames in mental models yet and explanations based on this line of thought are only speculative at this point in time.

Acceptance rates were very surprisingly low in Experiment 5, even compared to results in the literature that describe low acceptance rates in conditional reasoning tasks with natural material. Quinn & Markovits (1998) for example report rates of approx. 0.8 for MP, 0.65 for MT, and 0.5 for AC and DA. But the rates in Experiment 5 were far below these rates (0.44, 0.36, 0.27 and 0.28 for MP, MT, AC, and DA). The low acceptance rates are mostly due to the fact that participants accepted none of the inferences about 40 percent of the problems. This rejection pattern has not been reported in the literature nor has it been observed in Experiment 3 that was identical to Experiment 5 except for the explicit frequency information. Obviously, most participants did not follow the instruction to reason only from the major and minor premise. In the open comments participants could provide after the experiment, several participants expressed their confusion about the role of the explicit frequency information, they didn't know whether they were supposed to ignore it or not:

> *The task was unclear to me. Did I have to judge according to the given assumption or according to the given data* [meant is "explicit frequencies"]*? Why didn't the instruction mention the collection of data of 2000 alarms, rabbits, PC's?*
>
> *I didn't know whether I was supposed to go with the "Research Data"* [meant is "explicit frequencies"] *or with the thesis at the beginning of the questions.*[16]

There was no reference to the frequency information in the instruction, because the instruction in Experiment 5 was the one already used in Experiment 3. This was done to be able to directly compare the results from the two experiments. It was believed that the instruction was clear enough stressing that judgments should be based on the premises. But apparently the instruction was not transparent to all participants.

Since large effects of the frequency information were found in the acceptance rates of the inferences, it is plausible to assume that many participants considered the frequencies as relevant for the task. There were probably at least two ways to integrate the premises in solving the task: One solution is to understand the frequencies as part of the premises in a deductive way. These participants probably rejected any deductive inference, because a logically valid inference must be true under all circumstances and the presence of empirical counterexamples forbids drawing any inference with certainty. About a fourth of the participants shows this rejection pattern consistently over all four problems, about another third of participants in at least one of the problems. Comments expressing this solution were for example:

[16] Citations translated from German, originally:
"Die Fragestellung erschien mir unverständlich. Habe ich nun zu beurteilen, ob die Möglichkeiten anhand der gegebenen Annahme oder der gegebenen Daten zu beurteilen sind? Wieso wird in der vorhergehenden Anleitung die Datensammlung an 2000 Alarmen/Hasen/PCs nicht erwähnt?"
"Man weiß nicht ob man sich an den "Forschungsergebnissen" [meant is "explicit frequencies"] oder an der Eingangsthese orientieren soll!"

> *I always ticked NO, because the examples showed that all combinations exist – beetle with flower, beetle without flower, flower with beetle, flower without beetle etc. – so that strictly speaking nothing can be deduced, only with a certain probability.*
>
> *The task doesn't make sense according to logical deduction, because you never know a 100 percent. All combinations are possible and therefore no logical inference is possible.*[17]

Another option to integrate the frequencies in the task is to switch to an inductive judgment based on an probability judgment and to decide whether the inference is likely to be true. One comment clearly shows the conflict between deductive and inductive judgment:

> *Somehow, my answer is always NO, because I think that the inferences are never logically valid. Maybe I should have interpreted the task in percents and decide whether there is a high likelihood of the inference to be true?*[18]

Two aspects render it unlikely that the low acceptance rates of the inferences are due participants considering the major premise as false. First, among 394 participants there was not one expressing that idea in a comment. Second, in Experiment 4 only very few participants rejected the conditional statement as completely unbelievable by assigning it a degree of belief of zero.

In Experiment 5 the explicit frequency information clearly affected the acceptance rates of inferences in the way the probabilistic theory by Oaksford et al. (2000, Oaksford & Chater, 2001) predicts. These results fit in with results by Oaksford et al. (2000) who report effects of probabilities on the acceptability of inferences with abstract materials for which frequencies of all four relevant cases were explicitly given (Experiment 2) and with thematic materials pretested for the prior probabilities of *p* and *q* (Experiment 3). However, these results contradict recent findings by Oberauer, Weidenfeld, and Hörnig (2003) who found no effect of the prior probability of *p* and of *q* on the acceptance of MT and DA, and a small effect on AC in only one of two experiments. In Oberauer et al.'s experiments participants learned the probabilities in a probability learning task preceding the reasoning tasks, the learning success was controlled online during the task and again after the reasoning tasks which consisted of three blocks (Wason Selection Task, conditional reasoning tasks, and syllogisms). The instruction was a standard deductive instruction, very similar to the instruction used in Experiment 5. Oberauer et al.'s experiment and Experiment 5 differed in many aspects. Oberauer et al.'s experiments

---

[17] Citations translated from German, originally:
"[...] ich habe IMMER nein angeklickt, weil die beispiele ja immer zeigen, dass ALLE kombinationen auftreten - käfer mit blüte, käfer ohne blüte, blüte mit käfer,blüte ohne käfer etc. - so dass GAR nichts ganz streng gefolgert werden kann, nur mit irgendwelchen wahrscheinlichkeiten."
"das ergibt alles nach logischer schlussfolgerung keinen Sinn da man es nie 100 % sagen kann. denn es sind alle fälle offen und somit ist keine logische schlussfolgerung möglich."

[18] Citation translated from German, originally:
"Irgendwie ist meine Antwort jedesmal nein, da die Schlußfolgerungen meiner Meinung nach NIE logisch richtig sein müssen, hätte ich es vielleicht prozentual auswerten sollen, um aufeine hohe Wahrscheinlichkeit der Schlußfolgerung schließen zu können?"

used artificial materials, participants learned the probabilities in a separate phase prior to the reasoning tasks on a case-by-case basis, and participants learned the prior probabilities of *p* and of *q* (i.e., $P(p)$, $P(\neg p)$, $P(q)$, and $P(\neg q)$) while Experiment 5 used pseudo-natural material, the frequency information was given together with the cover story and the conditional reasoning task, and frequencies of conjunctions of events were given in a summary format. Experiment 5 used thematic contents, but frequency tables were assigned randomly to contents. Therefore, on average every content was presented equally often with every frequency table and the contents cannot be responsible for the observed effects of the frequency tables. Nonetheless, it cannot be excluded that the pseudo-natural nature of the conditionals causes the frequencies to be considered more relevant than in other contexts, e.g., than with abstract materials. A probably more crucial difference between the experiments is the sort of information participants received, prior probabilities versus frequencies of conjunctions of events. Frequencies of conjunctions have been shown in Experiment 4 to have a large effect on the believability of the major premise. However, it is unlikely that the observed effects of the frequencies on the acceptance rates of inferences are completely due to differences in the perceived degree of belief in the conditional. Because without further ad hoc assumptions this approach must assume that MP and MT are affected in the same way by the given frequencies – at least qualitatively. By the same token, this should be true for AC and DA. Essentially, the effects of the frequencies were comparable for MP and MT, but clearly different for AC relative to DA. But it's probably not the only difference between frequencies of conjunctions and prior probabilities that the former are known to affect the believability of the conditional. Moreover, participants might consider frequencies of conjunctions relevant for the tasks, but not prior probabilities per se. According to Oaksford et al. (2000) participants should base their response towards an inference on the conditional probability of the conclusion given the minor premise (e.g., $P(\neg p/\neg q)$ for MT). Normatively, these conditional probabilities increase with increasing prior probability of the conclusion (e.g., $P(\neg p)$ for MT). It has been shown that participants can compute conditional probabilities quite easily from conjunctive frequencies of cases (Oberauer & Wilhelm, 2003a). But it is not clear whether participants can use prior probabilities or differences in prior probabilities to estimate conditional probabilities. Maybe people regard prior probabilities as irrelevant if they don't have further knowledge of how the second feature is distributed within these prior probabilities of events. If the differences between Oberauer et al.'s experiments (2003a) and Experiment 5 are actually due to the fact that in one case prior probabilities have been presented and in the other case frequencies of conjunctions of events, this emphasizes once more the urge to develop a processural account of how people arrive at the estimates of conditional probabilities they use in reasoning tasks according to Oaksford et al. (2000). The currently normative account of Oaksford et al. cannot illuminate the observed discrepancies.

How did the presence versus absence of an explicitly mentioned causal link between antecedent and consequent affect the reasoning process respectively its output? Differences between causal and noncausal conditionals are not expected by the probabilistic approach to reasoning (Oaksford et al., 2000) or the modified mental model accounts (Markovits & Barrouillet, 2002; Schroyens & Schaeken, 2003). Nevertheless, participants were more willing to accept an inference when the story provided a causal link between antecedent and consequent. This was true for all inferences except MT. This observation replicates a similar finding by Valiña, Seoane, Ferraces, and Montserrat (1999). They compared reasoning from *natural probabilistic* causal conditionals (e.g., "If the miner smokes a lot, then he will get lung cancer") and noncausal conditionals (e.g., "If the sculptor cuts his hair, then he will get married") and found a large effect of causal structure: the acceptance rates for all four inferences were higher in the causal condition. The differences in acceptance rates might be due to differences in the degree of belief with which participants trust the major premises. For the MP and MT inferences, the presence respectively absence of a causal link furthermore mediated the ratio effect: in MP the ratio effect was stronger for the noncausal than the causal problems. In MT a ratio effect was found only for the noncausal problems, but not for the causal problems. The ratio was the factor of the frequency manipulation that largely affected MP and MT. The results are compatible with the idea that the frequency manipulation was more effective for the noncausal conditionals. If this line of thought is correct, it follows that the frequency effect for AC should be larger for the noncausal than for the causal problems. Descriptively, this was indeed the case although the trend did not reach significance, Wald $X^2(1) = 1.41$, $p = 0.23$, mean differences between acceptance rates between high and low frequency were 0.16 in the noncausal conditionals and 0.12 in the causal conditionals. In the result section of Experiment 4 (pp. 33-40) it was already suggested that for causal conditionals aspects aside from the explicit frequency information (e.g., believability or familiarity of the described causal relation or causal power of the causal mechanism) might play an important role in judging the believability of a conditional. The same argument can be made for the reasoning task in Experiment 5: it is possible that the frequency manipulation was more effective for the noncausal conditionals, because additional aspects might receive less weight in the noncausal than for the causal conditionals. Alternatively, the results can be interpreted as indicating that the presence of a causal link can compensate for a low ratio of *pq*-cases to *p¬q*-cases with respect to the acceptance of MP and MT.

4.3 Joint Analysis of Experiment 4 and 5

The integrated model developed at the beginning of this dissertation describes the interpretation and reasoning from causal conditionals - and partially noncausal conditionals as well. Is the model still valid if the objective probabilities are held constant through explicit frequency information as it has been done experimentally in Experiment 4 and 5? This question will be addressed in the following.

Experiment 4 and 5 did not measure all variables that were part of the original structural equation model (compare Figure 2, p. 20). There was no assessment of the subjective conditional probability, $P(q|p)$, and there was only one instruction group for the reasoning experiment. The model tested below includes the variables availability of exceptional situations, the degree of belief in a conditional, and the willingness to accept MP and the willingness to accept MT under a deductive instruction (compare Figure 12). For the path analysis data were aggregated over problems for 36 causal and 36 noncausal items. For each item mean values of degree of belief in the conditional, and mean acceptance rates of MP and MT were computed. Ratings of availability of exceptional situations were taken from Experiment 1 and 2, because these ratings are unbiased by the explicit frequency information added to the problems in Experiment 4 and 5. The rationale behind this technique is that frequency information should be irrelevant for the availability of exceptional situations.[19]

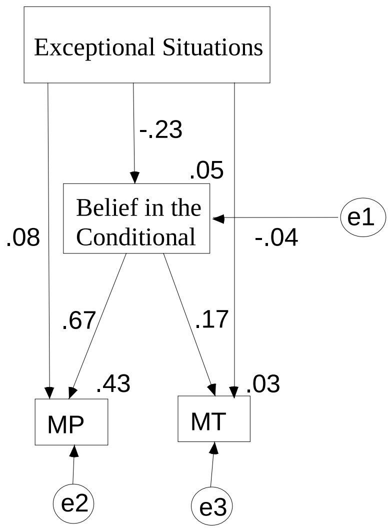


Figure 12: Path model with standardized path coefficients and R^2-values for the dependent variables

In the first step it was not distinguished between causal and noncausal problems. The path model depicted in Figure 12 was fitted to the overall data (N = 72). The model fit is very bad (X^2 = 18.4, df = 1; RMSEA = 0.50, range: 0.31-0.71; CFI = 0.70).[20] The degree of belief in the conditional enhanced the willingness to accept MP, this was the only path that received a significant path coefficient. Additionally, there was a trend that availability of exceptional situations decreased the degree of belief in a conditional, but this path coefficient was only marginally significant (p = 0.05). The standardized path coefficients are depicted in Figure 12.

The model fit was improved considerably by allowing the error terms of the willingness to accept MP and MT to covary. Furthermore, the direct pathway from the availability of exceptional situations to the acceptance of MP was fixed to be equal to the corresponding path for the acceptance of MT. The previous path analyses from Experiment 1 to 3 justify this restriction which gains one degree of freedom. The path model was fitted to the overall data, but this time different path coefficients were allowed for sub-models for the causal and noncausal conditionals alone, respectively. The standardized results can be seen in Figure 13. The model fits were satisfactory for the causal, but less sufficient for the noncausal conditionals as is evident from different measures of goodness of fits depicted in Table 15. The degree of belief in the conditional strongly affected the acceptance of MP, this path stands out in both sub-models (standardized path coefficients were 0.68 and 0.67 for the sub-model of the causal and noncausal problems alone). In the sub-model with the *causal* conditionals alone, clearly no other path received a significant weight (for details see Table 16). For the sub-model with the *noncausal* conditionals alone the degree of belief in the conditional positively affected furthermore the acceptance of MT (the

Table 15: Measures for Goodness of Fit

	χ^2	RMSEA	CFI
Overall	2.2 (df = 2)	0.04 (0-0.24)	0.997
Causal alone	0.0 (df = 1)	0.00 (0-0.22)	1.00
Noncausal alone	2.2 (df = 1)	0.18 (0-0.53)	0.963

[19] It should be noted that this procedure reduces the variability of the ratings of exceptional situations, because the same rating of availability of exceptional situations is assigned to four items sharing a content and a conditional, but with different frequency distributions.
[20] For the interpretation of the indices please compare footnote 12 at page 19.

path coefficient was 0.37, $p < 0.05$). Furthermore, availability of exceptional situations had a negative effect on the degree of belief in the conditional, although the path coefficient did not reach the conventional alpha level (the path coefficient was -0.31, $p = 0.06$).

Table 16: Standardized Path Coefficients and *p*-Values for Causal and Noncausal Problems

Path			Causal problems (n = 36)		Noncausal problems (n = 36)	
			Estimate	p	Estimate	p
Exceptional situations	-	Belief	0.00	0.99	-0.31*	0.06
Exceptional situations	-	MP	-0.10	0.40	0.11	0.32
Exceptional situations	-	MT	-0.09	0.40	0.15	0.32
Belief	-	MP	0.68*	0.00	0.67*	0.00
Belief	-	MT	0.01	0.97	0.37*	0.03

Note. Estimate = estimate of the standardized path coefficient. Belief = belief in the conditional. Exceptional situations = availability of exceptional situations. *$p < 0.10$.

Both sub-models explained 46 % of the variance of the willingness to accept MP (compare Figure 13). But the model did very badly at explaining the acceptance of MT in the causal condition ($R^2 = 0.01$) and only slightly better with the MT in the noncausal problems ($R^2 = 0.12$). The error terms of the acceptance of MP and MT covaried to a high degree (standardized values were 0.46 for causal and noncausal conditionals). This systematic covariation of MP and MT has to be generated by a factor (or several factors) that were not part of the path model. As has been shown above in the analysis of patterns of acceptance of inferences, in approximately 40 % of problems participants did not accept one single inference. This tendency to reject MP as well as MT appears to be the

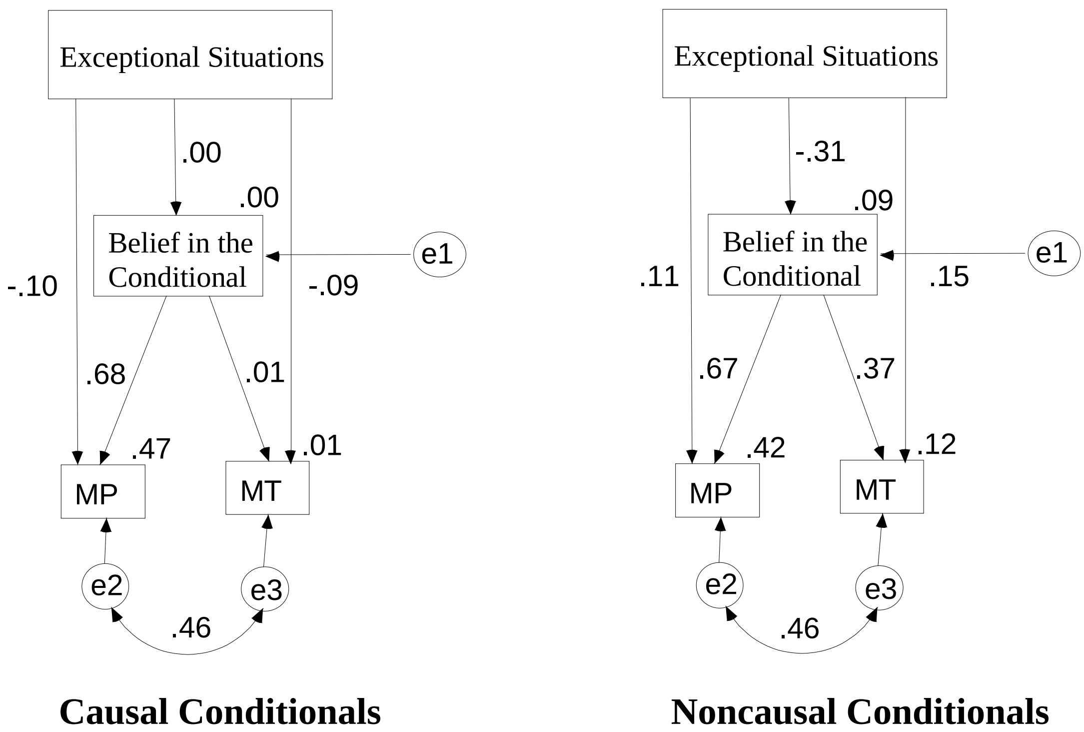


Figure 13: Structural equation models and standardized results. Please note that the rating of the availability of exceptional situations stems from Experiment 1 and 2, all other data are from Experiments 4 and 5.

most credible factor generating this common variance. Another plausible candidate is the logical status according to the material implication, because according to this standard both MP and MT are valid.

Discussion

In Experiments 1 to 3 the availability of exceptional situations suppressed the believability of the conditional. This suppression effect of exceptional situations on the degree of belief in the conditional was blocked (in the case of causal problems) respectively substantially reduced (in the case of noncausal problems) through the presence of explicit frequency information in Experiments 4 and 5. It seems remarkable that the availability of exceptions did not affect the degree of belief. A simple explanation can be provided if one assumes that participants use the availability of exceptions as a mean to estimate the degree of belief they put into a conditional statement. Frequency information renders this "detour" via exceptional situations unnecessary, if participants estimate the believability of the statement directly from the frequency information. But this explanation leaves the question open why this strategy should apply to the causal, but not completely to the noncausal problems. Which is especially puzzling if one bears in mind that other analyses of the data in Experiment 4 and 5 yielded hints that the frequency information was more effective for the noncausal than the causal conditionals. Despite the fact that the availability of exceptional situations did not affect the degree of belief, the degree of belief in a conditional still positively affected the willingness to accept MP and for the noncausal conditionals also the acceptance MT.

The path analyses showed furthermore that explicit frequency information blocked the suppression that exceptional situations directly exerted on MP and MT in Experiment 3, in both causal and noncausal problems. This finding is a surprise and more difficult to understand. The direct pathways rest on assumptions of the modified mental model accounts. In the discussion of the first three experiments (p. 24) two explanations were offered for the existence of the direct pathway simultaneously to the indirect pathway resting on probabilistic approaches. The results from Experiments 4 and 5 help to decide between these possible explanations. The first explanation argued that reasoning tasks differ from estimating the degree of belief in a conditional in such a way that the minor premise reactivates the whole knowledge structure that is associated with the problem and this additional activation is responsible for the direct effect of exceptional situations on the acceptance of MP and MT. There is no obvious reason why this should be different in problems that contain explicit frequency information. Therefore, this explanations seems implausible in the light of the new evidence.

The dual-process-specification suggested by Verschueren et al. (2003) provides an alternative explanation for the presence of an indirect coexisting with a direct pathway between exceptional situations and acceptance of MP and MT. According to this view, the direct pathway can be attributed to an analytic reasoning process that is very resource-consuming and that operates with availability of exceptional situations as input. But if frequency information is presented with the problems, it is plausible to assume that the frequencies can more easily serve as input for the analytic process than the availability of exceptional situations. Therefore, the direct pathway from availability of exceptional situations to MP and MT that represents the influence of the analytical process in Experiments 1 to 3 would be blocked by the presence of explicit frequency information in Experiment 5. The frequencies could be used by the analytical mechanism to directly assign weights to the model of counterexamples. This weight allocated to the counterexamples model would form the basis for accepting or rejecting the inference at hand. This criterion is in accordance with Schroyens et al. (2001) who argue that an inference is accepted if no counterexample is retrieved from memory or if a counterexamples is not very likely to be true. In the case of explicitly given frequencies of counterexamples, it would not matter whether a counterexample is retrieved from memory, but rather how high the likelihood of a counterexample is. Thus, the direct influence of exceptional situations would be overridden by the explicit frequency information. Frequency information would block the direct suppression of MP and MT through availability of exceptional situations.

5 Conclusions

The present dissertation focused on the question what meaning an everyday conditional statement conveys and what inferences it licenses. These questions were investigated with conditionals describing a fictitious causal relationship ("if cause, then effect", e.g., "If you fertilize a flower, then it will bloom") and with conditionals in which there was no meaningful relation between the propositions. It seemed advisable to confine the experiments to one semantic domain, to avoid getting lost in the "chameleon-like characteristics" of conditionals (Johnson-Laird & Byrne, 2002, p. 674).

Reasoning with Conditionals

At the beginning of the dissertation a model was developed to explain the understanding and reasoning from causal conditionals. The model integrates central assumptions from the conditional probability account of the meaning of conditionals (Edgington, 1995) and from two rival theories of conditional reasoning: the probabilistic approach (Oaksford et al., 2000) and modified mental model accounts (Markovits & Barrouillet, 2002; Schroyens & Schaeken, 2003). The model explains how the availability of exceptional situations reduces the willingness to accept MP and MT (e.g., a person is likely to infer from "If a match is struck, it lights" and "this match is struck", "therefore, it lights" unless she has in mind that *wet* matches do not light when they are struck). The model explained the data from Experiments 1 to 3 for the causal conditionals, but not for the arbitrary conditionals. Discrepancies between the causal and arbitrary conditionals in the path analyses were also observed in the data from Experiments 4 and 5. The discrepancies suggest that the underlying processes in understanding of and reasoning from meaningful conditionals might be fundamentally different from the corresponding processes in arbitrary conditionals. This emphasizes that a model developed and confirmed for one semantic domain need not be transferable to a different domain or type of domain. Surprisingly, the model for the causal conditionals yielded simultaneous support for the probabilistic approach and the modified mental model accounts to conditional reasoning. The probabilistic approach explains the suppression of MP and MT through the availability of exceptions by assuming that the availability of exceptions reduces the believability of the conditional (indirect pathway). The modified mental model accounts argue that participants perform an active search for exceptions whenever they are uncertain of the provisional conclusion (direct pathway). Obviously, both theories have their merits – at least in the domain of causal conditionals. How can the theories be reconciled? Following Verschueren et al. (2003) and Schroyens et al. (2003), the findings might be generated by two different reasoning mechanisms. This explanation is also consistent with further results from Experiments 4 and 5. Verschueren et al. (2003) assume a quick heuristic process yielding a degree of confidence in the putative conclusion as result by drawing on probability estimates directly retrieved from memory. The second mechanism is a slow analytical process that actively searches for counterexamples (i.e., exceptions) to the provisional conclusion. Both mechanisms operate within a critical time frame. If the analytical mechanism does not provide a counterexample within this time frame, the reasoner falls back to the output of the heuristic system and his decision will be based on the degree of confidence in the conclusion. A possible experiment testing this explanation straightforwardly is a design in which participants are asked to solve conditional reasoning tasks within a certain time frame. The time frame could additionally be varied between or within participants to realize different levels of time pressure. It would be expected that fast responses depend only on the degree of belief in a conditional and slow responses mainly on the availability of counterexamples – in accordance with findings of Verschueren et al. (2003). Thus, in the path model for the fast responses only the indirect pathway should receive significant weight. For the slow responses primarily the direct pathway is expected to obtain a significant path weight.

In Experiment 4 and 5 explicit frequency information of possibly relevant cases was added to the problems, four conditions with different frequency tables were implemented. This radically changed participants' answers to the deductive reasoning task. There were many participants who always or occasionally rejected all four inferences for a given problem. This complete rejection of inferences has not been reported in the literature nor was it observed in Experiment 3 where the same problems were used without explicit frequency information. The open commentaries at the end of the internet experiment indicate that some of the "rejecters" among the participants treated the frequencies as part of the premises. Following the deductive instruction they reached the decision that neither inference was necessarily true and could thus be drawn with certainty. Other participants also integrated the frequencies into their reasoning process, but instead of rejecting each inference produced a pattern that closely matched the predictions of the probabilistic theory of conditional reasoning by Oaksford et al. (2000). Some effects were moderated by the presence versus absence of a causal link between antecedent and consequent. MP and MT for example were less acceptable if the ratio of *pq*-cases to $p\neg q$-cases was low rather than high – as predicted by Oaksford et al. (2000). This effect was stronger for the arbitrary than the causal problems: The presence of a causal link was seemingly able to compensate for the low ratio. Since the relevant conditional probabilities were held constant between causal and arbitrary conditionals, Oaksford et al.'s theory cannot explain why and how the presence of a causal link should moderate the effects of the explicit frequency information. Further available evidence on the effects of probabilistic information on conditional reasoning tasks appears inconsistent: Probabilistic information affected reasoning outputs in some experiments (Oaksford et al., 2000), but not in others (Oberauer et al., 2003). The mentioned experiments differ from each other in manifold ways. Several factors seem especially important and could be used as starting points for further empirical work trying to resolve the discrepancies. One could suspect for example that it makes a difference whether the probabilistic knowledge is newly acquired in the experimental setting or is part of the background knowledge stored in long-term memory. If the knowledge is newly acquired in the experiment, the format with which participants are familiarized with the probabilities might play an important role, too. Formats previously used were probability learning tasks mimicking natural sampling procedures and provision of explicit probability/frequency information in a summary format. Another presumably relevant factor is what kind of probabilistic information is made accessible to the reasoner. Maybe participants can easily use frequencies/probabilities of conjunctions of events but not prior probabilities, although normatively prior probabilities are connected to the conditional probabilities relevant for the acceptance of MT, AC, and DA according to the account of Oaksford et al. (2000). The ultimate goal of this line of research would be the development of a detailed process theory that explains when and how participants are able to and do use probabilistic information in (conditional) reasoning. The current approach of Oaksford and colleagues (Oaksford et al., 2000; Oaksford & Chater, 2001) is a normative theory predicting how people should behave rationally, but contains no process assumptions specifying how people (try to) arrive at these optimal solutions. Whether this belittles successful predictions derived by the theory depends on the standpoint of the observer, but the lack of a process theory is clearly unsatisfactory. Further empirical and theoretical work seems strongly desirable.

Understanding conditionals

What is the meaning conveyed by an ordinary indicative conditional? According to the conditional probability approach of the meaning of conditionals (Edgington, 1991, 1995, 2003; Oberauer & Wilhelm, 2003a), a conditional communicates that the consequent is likely given the antecedent. The sentence "If you fertilize a flower, then it will bloom" hence means that a fertilized flower has a high probability of blooming. The experiments reported here strongly supported the conditional probability approach. In Experiment 4 for example, approximately 40 percent of the participants were classified as following the conditional probability approach. The behavior of a smaller group of participants (approx. 20 %) agreed with predictions made by the mental model theory assuming that participants don't flesh out the initial model of the conditional. For the arbitrary conditionals there was also some support for a mental model interpretation according to which participants construct two mental models to represent a biconditional interpretation of the conditional. According to the traditional mental model theory (Johnson-Laird, 1983) conditional statements constitute truth-functional assertations. In the mental model approach without fleshing-out for example the *pq* case is considered as true, the rest of the cases of the truth table are disregarded unless the model is fleshed out. A biconditional interpretation respects additionally the $\neg p \neg q$-case as true. There is no combination of truth values that embodies the conditional probability of the consequent given the antecedent. Hence, the conditional probability approach views conditionals as not truth-functional. In accordance with this distinction, the two groups of participants who have been identified in Experiment 4 and previously by Oberauer and Wilhelm (2003a) as well as Evans et al. (2003) can be described as operating with a probabilistic and with a truth-functional interpretation of conditionals respectively. The question arises whether a person constantly follows either a probabilistic or a truth-functional interpretation or whether the same person can switch between interpretations. And if interpretations vary within a person, is it possible to identify the factors that either trigger a probabilistic or a truth functional reading? These empirical questions are yet unapproached. But before examining individual or task-dependent differences between a probabilistic and a truth-functional interpretation, it should be made sure that there actually is a truth-functional interpretation at all. Support for the mental model theory without elaboration stemmed from the finding that some participants assigned lower degrees of belief to conditionals which were associated with few cases of *pq* as opposed to many cases of *pq*. This frequency of *pq* effect is compatible with the mental model theory without elaboration, but it is likewise consistent with a probabilistic interpretation in a biconditional reading. A biconditional reading can be expressed as "if *p* then *q*, and if *q* then *p*". If participants with a probabilistic interpretation of conditionals represent and consider additionally the reverse conditional probability, $P(p|q)$, a frequency of *pq* effect would result. To be able to correctly attribute the frequency of *pq* effect to a truth functional or a probabilistic interpretation, an experiment would be needed that manipulates the reverse conditional probability while holding constant the frequency of *pq* within each level of reverse conditional probability. In this experiment, it would furthermore be interesting to examine what kind of conditional probability is relevant for judging backward causal conditionals in which the antecedent describes the effect and the consequent the cause (i.e., "if effect, then cause"). The probabilistic approach to the meaning of conditionals regards the conditional probability of the consequent given the antecedent, $P(q|p)$, as the major component for the acceptance of any conditional statement, hence also for these conditionals, whereas theories of causal induction (e.g., by Cheng, 1997) identify the probability of the effect given the cause, *P*(effect|cause), as the most relevant factor, a probability which equals in this case the reverse conditional probability, $P(p|q)$.

In addition, after the experiment participants could be asked to describe what frequencies they consider relevant for the task of assessing the believability of the conditionals and whether they are aware of how they arrived at their estimates. Previous experiences with self-reported strategy use are encouraging, but nonetheless advise caution in interpreting the self-reports: In contingency judgment tasks for example, Anderson and Sheu (1995) found that participants specified quite accurately how they arrived at their judgments, whereas in Mandel and Lehman's (1998) experiments, the strategy reports were only loosely associated with actual individual weightings of cells in the contingency table.

Varying the probability of the consequent given the antecedent, $P(q|p)$, affected causal and arbitrary conditionals in the same way: Higher degrees of belief were assigned if the conditional probability was high, a finding predicted by the probabilistic approach to the meaning of conditionals by Edgington. Theories of causal induction consider not only $P(q|p)$ as crucial, but also the probability of the consequent given the *absence* of the antecedent, $P(q|\neg p)$. The perceived causal strength with which a cause generates its effect for examples depends in the causal power theory of Cheng (Cheng, 1997; Cheng & Novick, 1992) on the probabilistic contrast that is defined as:

$$P(\text{effect} \mid \text{cause}) - P(\text{effect} \mid \neg\text{cause})$$

If a gardener wants to know for example whether a fertilizer is effective in making flowers bloom, he has to look how many of the fertilized flowers bloom (i.e., P(effect|cause)) and concurrently how many non-fertilized flowers bloom (i.e., P (effect|¬cause)).[21] A fertilizer can only be regarded as effectual if the first rate is substantially higher than the second one. Otherwise, the gardener could spend his money more wisely. Do noncausal conditionals also depend on both conditional probabilities? In Experiments 4 and 5 $P(q|\neg p)$ was held constant at 0.5 to avoid confounds. Further experiments varying $P(q|p)$ and $P(q|\neg p)$ independently from each other, could resolve whether people integrate both conditional probabilities if asked for a believability estimate in the way predicted by the causal power theory for judgments of the causal power. This was recently also suggested by Over and Evans (2003), but not tested empirically. Empirical evidence that people's evaluations of conditionals depend on frequencies of $\neg p$-cases would come as a real surprise. Empirical findings in the literature demonstrate that in several tasks trying to assess the meaning people connect to a conditional, cases with false antecedents are treated as irrelevant. For example, Johnson-Laird and Tagart (1969), showed participants pictures corresponding to the four cases of a conditional's truth table (pq, $p\neg q$, $\neg pq$, $\neg p\neg q$) and asked them to decide whether the conditional is "true", "false" or "irrelevant" given the state of affairs on the picture. The conditional was considered to be "irrelevant" for pictures with false antecedents. Likewise, when asked to generate examples of cases that either make a given conditional true or false, participants mainly produced pq-cases or $p\neg q$-cases, respectively, but only rarely cases with false antecedents (Evans, 1972). Similarly, in contingency judgment tasks the cells corresponding to the false antecedent cases are largely neglected (Over & Green, 2001; Mandel & Lehman, 1998).

Over and Green (2001) argue that it is adaptive to put more weight on the first term in the causal contrast (i.e., $P(q|p)$ or P(effect|cause)) if causes and effects are rare. But if causes and effects are frequent, people should weight the second term in the causal contrast more strongly (i.e., $P(q|\neg p)$ or P (effect|¬cause)). If for example most of the gardener's flowers are fertilized and most of them bloom, the few negative instances that have not been fertilized are very informative: If the non-fertilized flowers bloom as well, the fertilizer is probably not very effective. To test this hypothesis experimentally, experiments are needed manipulating the prior probabilities of p and q, $P(p)$ and $P(q)$, systematically and as much as possible independently from $P(q|p)$ and $P(q|\neg p)$.[22]

The conditional probability approach predicts that subjective conditional probability of the consequent given the antecedent and degree of belief go always hand in hand. Experiments presented here tested the approach correlationally with pseudo-natural materials naturally varying in believability and by manipulation of the objective conditional probabilities. Further empirical tests should employ other methods as well. For example, the context could be manipulated in such a way that the conditional is uttered either by an expert or a lay person thus manipulating the source confidence. This paradigm was for example successfully used by Stevenson and Over (2001). Any dissociation between the subjective conditional probability and the degree of belief in a conditional represents a challenge for Edgington's approach. There are causal relations that seem to suggest such a dissociation and might constitute interesting test cases for the

[21] This corresponds to the control group in an experimental design.

[22] In a two-by-two contingency table, $P(p)$, $P(q)$, $P(q|p)$ and $P(q|p)$ cannot vary completely independently from each other.

theory. "Smoking causes cancer" for example is undoubtedly a highly believable statement since it seems to be part of the shared general knowledge in our society, but the corresponding conditional probability that somebody who smokes actually has cancer, is probably perceived as low. Can such causal relations be meaningfully translated into conditional statements? If so, how do participants evaluate the corresponding conditional probabilities and degrees of belief?

The Role of a Causal Link for the Believability of a Conditional

Philosophers have suggested to consider conditionals in which antecedent and consequent are irrelevant for each other like "If Napoleon is dead, then Bristol is in England" as not acceptable or even false. Similarly, Edgington (1995) argues that these kind of conditionals are misleading for the addressee. These positions seem to claim the presence of a meaningful connection between the propositions in a conditional as a necessary condition for its eligibility of existence. This extreme viewpoint can be contrasted with the hypothesis that the presence of a meaningful relation between antecedent and consequent enhances a conditional's believability (semantic link hypothesis). In this dissertation, a paradigm was developed to compare conditionals with and without a meaningful relation between the propositions. The paradigm controlled for a variety of context and content effects. To achieve this goal identical wordings of conditionals were used in both conditions. The very same conditional statement was either embedded in a cover story explicitly mentioning a causal link between antecedent and consequent (causal conditionals) or in a neutral cover story providing no meaningful relation between the two propositions (noncausal or arbitrary conditionals). Unlike philosophers who deem arbitrary conditionals as unacceptable or false, participants in Experiments 1, 2, and 4 did not reject arbitrary conditionals per se as unbelievable. Nonetheless, arbitrary conditionals were judged to be less believable than causal ones. This effect persisted even when the conditional probabilities of the consequent given the antecedent were held constant experimentally in Experiment 4. Thus, there was no support for the idea that a meaningful relation constitutes a necessary condition for a true or believable conditional. But the mere presence of a causal relation between the propositions awarded additional credibility to the conditional. This finding is consistent with the semantic link hypothesis, but also with the conditional probability account if one suspects that the experiment controlled the objective conditional probabilities, but not necessarily the subjective conditional probabilities. To decide between the two explanations, one would need a supplementary experiment assessing estimates of the relevant subjective conditional probability, $P(q|p)$, for the items used in Experiment 4 and 5. An especially interesting outcome would be if the presence of a causal link did not affect the subjective conditional probability, but the degree of belief in the conditional. With this pattern of results the interpretation would seem warranted that the presence of a causal link is a relevant factor contributing to the believability of a conditional – independently of the objective and subjective conditional probability. This would strongly support the semantic link hypothesis and challenge the view that conditionals can be completely reduced to conditional probabilities as suggested by the probabilistic approach to the meaning of conditionals by Edgington.

The word " if " will probably remain intuitively obvious for the layman and puzzling for the expert for a long time to come. But it has been shown that recent probabilistic approaches bear good prospects for successfully identifying the psychological processes underlying the everyday understanding of and the reasoning with conditionals.

References

Anderson, J. R., & Sheu, C. F. (1995). Causal inferences as perceptual judgments. *Memory and Cognition*, *23*, 510-524.

Barrouillet, P. (1996). Transitive inferences from set-inclusion relations and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1408-1422.

Barrouillet, P., Grosset, N., & Lecas, J.-F. (2000). Conditional reasoning by mental models: Chronometric and developmental evidence. *Cognition*, *74*, 1-30.

Barrouillet, P., & Lecas, J.-F. (1998). How can mental models theory account for content effects in conditional reasoning? A developmental perspective. *Cognition*, *67*, 209-253.

Bindra, D., Clarke, K. A, & Shultz, T. R. (1980). Understanding predictive relations of necessity and sufficiency in formally equivalent "causal" and "logical" problems. *Journal of Experimental Psychology: General*, *4*, 422-443.

Byrne, R. M. (1989). Suppressing valid inferences with conditionals. *Cognition*, *31*, 61-83.

Byrne, R. M., & Walsh, C. R. (2002).Contradictions and counterfactuals: Generating belief revisions in conditional inference. *Proceedings of the 24rd Annual Meeting of the Cognitive Science Society, Fairfax.*

Chan, D., & Chua, F. (1994). Suppression of valid inferences: Syntactic views, mental models, and relative salience. *Cognition*, *53*, 217-238.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367-405.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391-416.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*, 365-382.

Cummins, D. D. (1995). Naive theories and causal deduction. *Memory and Cognition*, *23*, 646-658.

Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory and Cognition*, *19*, 274-282.

Dieussaert, K., Schaeken, W., & d'Ydewalle, G. (2002). The relative contribution of content and context factors on the interpretation of conditionals. *Experimental Psychology*, *49*, 181-195.

Edgington, D. (1991). Do conditionals have truth conditions? In Jackson, F. (Ed.), *Conditionals* (pp. 176-201). Oxford: Oxford University Press.

Edgington, D. (1995). On Conditionals. *Mind*, *104*, 235-329.

Edgington, D. (2003). What if? Questions about conditionals. *Mind and Language*, *18*, 380-401.

Evans, J. St. B. T. (1972). Interpretation and matching bias in a reasoning task. *Quarterly Journal of Experimental Psychology, 24,* 193-199.

Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*, 321-335.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning.* Hove: Psychology Press.

Fairley, N., Manktelow, K., & Over, D. (1999). Necessity, sufficiency, and perspective effects in causal conditional reasoning. *Quarterly Journal of Experimental Psychology*, *52A*, 771-790.

Fillenbaum, S. (1977). Mind your p' and q's: The role of content and context in some uses of and, or, and if. In G.H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 11, pp. 411-00). New York: Academic Press.

Gavanski, I., & Hui, C. (1992). Natural sample spaces and uncertain belief. *Journal of Personality and Social Psychology*, *63*, 766-780.

Geis, M. L., & Zwicky, A. M. (1971). On invited inferences. *Linguistic Inquiry, 2,* 561-566.

George, C. (1995). The endorsement of the premises: Assumption-based or belief-based reasoning. *British Journal of Psychology*, *86*, 93-111.

George, C. (1997). Reasoning from uncertain premises. *Thinking and Reasoning*, *3*, 161-189.

George, C. (1999). Evaluation of the plausibility of a conclusion derivable from several arguments with uncertain premises. *Thinking and Reasoning*, *1999*, 245-281.

Grice, H. P. (1975). Logic and converstaion. In P. Cole & J. L. Morgan (Eds), *Syntax and semantics: Vol. 3 Speech acts* (pp 41-48). New York: Academic Press.

Hadjichristidis, C., Stevenson, R. J., Over, D. E., Sloman, S. A., Evans, J. St. B. T., & Feeney, A. (2001). On the evaluation of "if p then q"-conditionals. *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society, Edinburgh.*

Hanita, M., Gavanski, I., & Fazio, R. H. (1997). Influencing probability judgments by manipulating the accessibility of sample spaces. *Personality and Social Psychology Bulletin*, *23*, 801-813.

Johnson-Laird, P. N. (1983). *Mental Models.* Cambridge: University Press.

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences*, *5*, 434-442.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646-678.

Johnson Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J. P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, *106*, 62-88.

Johnson-Laird, P. N., & Tagart, J. (1969). How implication is understood. *American Journal of Psychology, 82,* 367-373.

Liu, I. M. (2003). Conditional reasoning and conditionalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 694-709.

Liu, I., Wu, J., & Lo, K. (1996). A probabilistic interpretation of "If-Then". *Quarterly Journal of Experimental Psychology*, *49A*, 828-844.

Mandel, D. R., & Lehman, D. R. (1998). Integration of contingency information in judgments of cause, covariation, and probability. *Journal of Experimental Psychology: General*, *127*, 269-285.

Marcus, S. L., & Rips, L. J. (1979). Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior*, *18*, 199-223.

Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review*, *22*, 5-36.

Markovits, H., &. Doyon, C., & Simoneau, M. (2002). Individual Differences in working memory and conditional reasoning with concrete and abstract content. *Thinking and Reasoning*, *8*, 97-107.

Markovits, H., Fleury, M. L., Quinn, S., & Venet, M. (1998). The development of conditional reasoning and the structure of semantic memory. *Child Development*, *69*, 742-755.

Markovits, H., & Potvin, F. (2001). Suppression of valid inferences and knowledge structures: The curious effect of producing alternative antecedents on reasoning with causal conditionals. *Memory and Cognition, 29*, 736-744.

Markovits, H., & Quinn, S. (2002). Efficiency of retrieval correlates with "logical" reasoning from causal conditional premises. *Memory and Cognition*, *30*, 696-706.

McMullen, M. N., Fazio, R. H., & Gavanski, I. (1997). Motivation, attention, and judgment: A natural sample spaces account. *Social Cognition*, *15*, 77-90.

Newstead, S. E., Ellis, M. C., Evans, J. St. B. T., & Dennis, I. (1997). Conditional reasoning with realistic material. *Thinking and Reasoning*, *3*, 49-76.

O'Brien, D. P., Costa, G., & Overton, W. F. (1986). Evaluations of causal and conditional hypotheses. *Quarterly Journal of Experimental Psychology*, *38A*, 493-512.

Oaksford, M., & Chater, N. (2001). The probabilistic approach to human reasoning. *Trends in Cognitive Sciences*, *5*, 349-357.

Oaksford, M., & Chater, N. (2003a). Computational levels and conditional inference: Reply to Schroyens and Schaeken (2003). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 150-156.

Oaksford, M., & Chater, N. (2003b). Conditional probability and the cognitive science of conditional reasoning. *Mind and Language*, *18*, 359-379.

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 883-899.

Oberauer, K., Weidenfeld, A., & Hörnig, R. (2003). Logical reasoning and probabilities – a comprehensive test of Oaksford and Chater (2001). (Unpublished manuscript).

Oberauer, K., & Wilhelm, O. (2003a). The meaning(s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 680-693.

Oberauer, K., & Wilhelm, O. (2003b). The meaning(s) of "if": Conditional probabilities and mental models. *Proceedings of the 25rd Annual Meeting of the Cognitive Science Society, Boston.*

Over, D. E., & Evans, J. St. B. T. (2003). The probability of conditionals: The psychological evidence. *Mind and Language*, *18*, 340-358.

Over, D. E, & Green, D. W. (2001). Contingency, causation, and adaptive inference. *Psychological Review*, *108*, 682-684.

Peel, E. A. (1967). A method for investigating children's understanding of certain logical connectives used in binary propositional thinking. *British Journal of Mathematical and Statistical Psychology*, *20*, 81-92.

Politzer, G., & Bourmaud, G. (2002). Deductive reasoning from uncertain conditionals. *British Journal of Psychology*, *93*, 345-381.

Quinn, S., & Markovits, H. (1998). Conditional reasoning, causality, and the structure of semantic memory: Strength of association as a predictive factor for content effects. *Cognition*, *68*, B93-B101.

Reips, U. D. (2002). Standards for internet-based experimenting. *Experimental Psychology*, *49*, 243-256.

Rijmen, F., & De Boeck, P. (2003). A latent class model for individual differences in the interpretation of conditionals. *Psychological Research*, *67*, 219-231.

Rips, L. J. (2001). Two kinds of reasoning. *Psychological Science*, *12*, 129-134.

Schroyens, W., & Schaeken, W. (2003). A critique of Oaksford, Chater, and Larkin's (2000) conditional probability model of conditional reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 140-149.

Schroyens, W., Schaeken, W., & d'Ydewalle, G. (2001). The processing of negations in conditional reasoning: A meta-analytic case study in mental model and/or mental logic theory. *Thinking and Reasoning*, *7*, 121-172.

Schroyens, W., Schaeken, W., & Handley, S. (2003). In search of counter-examples: Deductive rationality in human reasoning. *The Quarterly Journal of Experimental Psychology*, *56A*, 1129-1145.

Sherman, S. J., McMullen, M. N., & Gavanski, I. (1992). Natural sample spaces and the inversion of conditional judgments. *Journal of Experimental Social Psychology*, *28*, 401-421.

Sperber, D., & Wilson, D. (1995). *Relevance: communication and cognition* (2nd edition). Malden: Blackwell.

Stevenson, R. J., & Over, D. E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology, 48A*, 613-643.

Stevenson, R. J., & Over, D. E. (2001). Reasoning from uncertain premises: Effects of expertise and conversational context. *Thinking and Reasoning*, *7*, 367-390.

Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory and Cognition*, *22*, 742-758.

Thompson, V. A. (1995). Conditional reasoning: The necessary and sufficient conditions. *Canadian Journal of Experimental Psychology*, *49*, 1-60.

Thompson, V. A. (1996). Reasoning from false premises: The role of soundness in making logical deductions. *Canadian Journal of Experimental Psychology*, *50*, 315-319.

Thompson, V. A., & Mann, J. M. (1995). Perceived necessity explains the dissociation between logic and meaning: The case of "Only If". *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1554-1567.

Torrens, D., Cramer, K. M., & Thompson, V. A. (1999). Individual differences and the belief bias effect: Mental models, logical necessity, and abstract reasoning. *Thinking and Reasoning*, *5*, 1-28.

Traxler, M. J., Aked, J. P., Moxey, L. M., & Sanford, A. J. (1997). Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 88-101.

Tversky, A., & Kahneman, D. (1980). Causal Schemas in Judgment under Uncertainty. In M. Fishbein (Ed.), *Progress in Social Psychology* (pp. 49-72). Hillsdale NJ: Erlbaum.

Vadeboncoeur, I., & Markovits, H. (1999). The effect of instructions and information retrieval on accepting the premises in a conditional reasoning task. *Thinking and Reasoning*, *5*, 97-113.

Valiña, D., Seoane, G., Ferraces, J., & Martin, M. (1999). The importance of pragmatic aspects in conditional reasoning. *Spanish Journal of Psychology*, *20*, 20-31.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2003). A dual process specification of causal conditional reasoning. (Unpublished manuscript).

Waldmann, M. R. (2001). Predictive versus diagnostic learning: evidence from an overshadowing Paradigm. *Psychonomic Bulletin and Review*, *8*, 600-608.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, *121*, 222-236.

Appendix

Appendix A: Original Materials

List of Original Materials Used in Experiment 1 to 5 (Cover Stories and Conditionals)

1. Allergic Disease

Wenn ein Hund unter Midosis leidet,
dann läßt sich Xathylen in seinem Blut nachweisen

I. Causal Forward
In einem großen Labor für Tiermedizin in Australien ist vor kurzem eine neue allergische Krankheit bei Hunden entdeckt worden. Die Forscher haben die Krankheit Midosis getauft. Inzwischen wissen die Wissenschaftler schon viele Details über Midosis. Midosis führt unter anderem dazu, daß im Blut der betroffenen Tiere die bisher unbekannte Substanz Xathylen gebildet wird.

II. Causal Backward
In einem großen Labor für Tiermedizin in Australien ist vor kurzem eine neue allergische Krankheit bei Hunden entdeckt worden. Die Forscher haben die Krankheit Midosis getauft. Inzwischen wissen die Wissenschaftler schon viele Details über Midosis. Im Blut der betroffenen Tiere entsteht eine bisher unbekannte Substanz namens Xathylen, diese Substanz führt zu den vielfältigen Symptomen der Midosis.

III. Noncausal
Die Forschungsbereiche eines großen Labors für Tiermedizin in Australien sind physiologische Grundlagenforschung (z.B. die Zusammensetzung von Blut bei verschiedenen Tierarten) und allergische Krankheiten. Das Labor hat vor kurzer Zeit eine neue allergische Krankheit bei Hunden namens Midosis entdeckt. Eine andere Abteilung hat eine bisher unbekannte Substanz im Blut von Katzen entdeckt und sie Xathylen getauft. In den letzten Wochen haben die Forscher erforscht, ob Xathylen auch im Blut von Hunden vorkommt.

2. Tropical Plant

Wenn eine Pherotelia blüht,
findet man auf der Pflanze Blaupunktkäfer

I. Causal Forward
Biologen einer amerikanischen Universität haben vor kurzer Zeit eine neue tropische Pflanze entdeckt und sie Pherotelia getauft. Blüht die Pherotelia, lockt ihr Pollen

Blaupunktkäfer an. Diese versammeln sich auf ihren Blättern, denn die Blaupunktkäfer ernähren sich vom Pollen der Blüte.

II. Causal Backward
Biologen einer amerikanischen Universität haben vor kurzer Zeit eine neue tropische Pflanze entdeckt und sie Pherotelia getauft. Eine besondere Käferart, die Blaupunktkäfer, bringen die Pherotelia zum Blühen, indem sie ihre Eier in ihre Blätter und Stiele legen.

III. Noncausal
Biologen einer amerikanischen Universität haben vor kurzer Zeit eine neue tropische Pflanze entdeckt und sie Pherotelia getauft. Sie blüht zweimal im Jahr. Sie wächst nur in Äquatornähe und bevorzugt nährstoffarme Böden. In dem Gebiet, wo die Pherotelia wächst, lebt ebenfalls der seltene Blaupunktkäfer.

3. Mechanical Object of Art

Wenn das Licht an ist,
hört man das Lied

I. Causal Forward
Susanne ist Künstlerin, sie erstellt auf Bestellung auch sehr ungewöhnliche figurale Objekte. Für Daniel, der von Beruf Ingenieur ist, hat sie in den letzten Wochen eine spezielle Figur geschaffen. Es handelt sich um ein Objekt aus Stahl von Größe und ungefährer Form einer Schuhschachtel. Schüttelt man das Objekt, geht innen eine helle Lampe an. Ein Lichtmesser reagiert darauf und aktiviert einen kleinen Lautsprecher, aus dem daraufhin ein Lied ertönt.

II. Causal Backward
Susanne ist Künstlerin, sie erstellt auf Bestellung auch sehr ungewöhnliche figurale Objekte. Für Daniel, der von Beruf Ingenieur ist, hat sie in den letzten Wochen eine spezielle Figur geschaffen. Es handelt sich um ein Objekt aus Stahl von Größe und ungefährer Form einer Schuhschachtel. Schüttelt man das Objekt, ertönt ein Lied und ein Schallmeßgerät reagiert darauf und veranlaßt, daß innen eine helle Lampe angeht.

III. Noncausal
Susanne ist Künstlerin, sie erstellt auf Bestellung auch sehr ungewöhnliche figurale Objekte. Für Daniel, der von Beruf Ingenieur ist, hat sie in den letzten Wochen eine spezielle Figur geschaffen. Es handelt sich um ein Objekt aus Stahl von Größe und ungefährer Form einer Schuhschachtel. Manchmal ertönt daraus ein Lied und manchmal brennt innen eine helle Lampe.

4. DNA-Mutation

**Wenn bei einem Kaninchen Natrolsan nachgewiesen wird,
findet man die DNA-Mutation**

I. Causal Forward
In einer abgelegenen chinesischen Provinz hat die Chemiefirma "Science United" in den letzten Jahren ein neues Pflanzenschutzmittel getestet. Es enthält u.a. die neue Substanz Natrolsan. Die Forscher, die den Versuch überwachen, haben jetzt festgestellt, daß über die Nahrung aufgenommenes Natrolsan bei Kaninchen eine bestimmte neue DNA-Mutation auslöst.

II. Causal Backward
In einer abgelegenen chinesischen Provinz hat die Chemiefirma "Science United" in den letzten Jahren ein neues Pflanzenschutzmittel getestet. Die Forscher, die den Versuch überwachen, haben jetzt festgestellt, daß das Pflanzenschutzmittel bei den einheimischen Kaninchen eine bestimmte neue DNA-Mutation auslöst. Diese Mutation führt dazu, daß im Körper des Kaninchens eine neue Substanz namens Natrolsan gebildet wird.

III. Noncausal
In einer abgelegenen chinesischen Provinz erforschen westliche Biologen der Chemiefirma "Science United" die einheimischen Säugetiere. Eine Teilgruppe der Biologen interessiert sich besonderes für eine bestimmte neue DNA-Mutation bei Kaninchen. Einer anderer aus der Gruppe ist Biochemiker und forscht über eine neulich entdeckte Substanz namens Natrolsan, die im Körper von manchen Nagetieren vorkommt.

5. Tribal Behavior

**Wenn ein Mann Zenobia-Kraut raucht,
leidet er unter Haarausfall**

I. Causal Forward
Maria ist Entwicklungshelferin in Südamerika. Sie betreut als Ärztin das Gebiet der Yamarati-Indianer. Viele Männer dieses Stammes rauchen getrocknetes Zenobia-Kraut. Zum Leidwesen der Männer verursacht es Haarausfall.

II. Causal Backward
Maria ist Entwicklungshelferin in Südamerika. Sie betreut als Ärztin das Gebiet der Yamarati-Indianer. Viele Männer des Stammes rauchen getrocknetes Zenobia-Kraut, denn die Yamarati sind überzeugt, daß es gegen Haarausfall hilft.

III. Noncausal
Maria ist Entwicklungshelferin in Südamerika. Sie betreut als Ärztin das Gebiet der Yamarati-Indianer. Sie macht sich Sorgen, weil sehr viele Männer des Stammes Zenobia-Kraut rauchen, dessen gesundheitsschädigende Wirkung der von Tabak gleicht. Ein Kollege aus Deutschland, der sich für die Häufigkeit von Haarausfall in unterschiedlichen Kulturen interessiert, hat sie neulich gebeten, für ihn in ihren Krankenakten zusätzlich zu notieren, welche der Patienten unter Haarausfall leiden.

6. Computer Virus

Wenn das Datum des PC's auf dem 24.12.01 steht,
ist die Startseite des Internetbrowsers www.witz.com

I. Causal Forward
Vor einigen Tagen haben die Sicherheitsfachleute der Agentur "Fun Media" entdeckt, daß ein Hacker einen Virus in ihr internes Netz eingeschleust hat. Der Virus verbreitet sich über das E-Mail-Programm. Sie wissen bereits, daß er dadurch aktiviert wird, daß das Datum des Computers auf den 24.12.01 springt. Die Fachleute versuchen gerade herauszufinden, welche Schäden der aktivierte Virus dann anrichtet. Sie haben bisher entdeckt, daß der Virus die Startseite des Internetbrowers auf die Internetseite www.witz.com umstellt.

II. Causal Backward
Vor einigen Tagen haben die Sicherheitsfachleute der Agentur "Fun Media" entdeckt, daß ein Hacker einen Virus in ihr internes Netz eingeschleust hat. Der Virus verbreitet sich über das E-Mail-Programm. In der zugehörigen E-Mail wird zum Besuch der Internetseite www.witz.com aufgefordert. Wird diese tatsächlich besucht, wird dadurch der Virus aktiviert. Die Fachleute versuchen gerade herauszufinden, welche Schäden der aktivierte Virus dann anrichtet. Sie haben bisher entdeckt, daß er das Datum des betroffenen Computers auf den 24.12.01 setzt.

III. Noncausal
Vor einigen Tagen haben die Sicherheitsfachleute der Agentur "Fun Media" entdeckt, daß ein Hacker einen Virus in ihr internes Netz eingeschleust hat. Der Virus verbreitet sich über das E-Mail-Programm. Die Fachleute untersuchen nun alle Computer auf den Virus. Als Dokumentation notieren sie die Kennziffer des PC’s, seine jeweilige Systemzeit, das Systemdatum, die Startseite des Internetbrowers, das verwendete E-Mail-Programm, die Kennziffer der Office-Version und natürlich, ob der Computer mit dem Virus infiziert war.

7. Social Relationship

**Wenn Katrin aggressiv nachfragt,
zieht Michael sich in Schweigen zurück**

I. Causal Forward
Katrin und Michael machen seit einem Jahr eine Paartherapie. Die Therapeutin hat einige Gesetzmäßigkeiten herausgefunden, nach denen ihre Beziehung und ihr Verhalten im Streit funktioniert. Sie hat zum Beispiel festgestellt, daß Katrins aggressives Fragen nach den Ursachen des Konflikts in Konfliktsituationen regelmäßig dazu führt, daß Michael sich in totales Schweigen zurückzieht.

II. Causal Backward
Katrin und Michael machen seit einem Jahr eine Paartherapie. Die Therapeutin hat einige Gesetzmäßigkeiten herausgefunden, nach denen ihre Beziehung und ihr Verhalten im Streit funktioniert. Sie hat zum Beispiel festgestellt, daß Michael sich in Konfliktsituationen oft in totales Schweigen zurückzieht. Dies löst bei Katrin aggressives Nachfragen nach den Ursachen des Konfliktes aus.

III. Noncausal
Katrin und Michael machen seit einem Jahr eine Paartherapie. Die Therapeutin hat einige Gesetzmäßigkeiten herausgefunden, nach denen ihre Beziehung und ihr Verhalten im Streit funktioniert. In Konfliktsituationen zeigen die beiden unterschiedliche Verhaltensweisen. Bei Michael beobachtet die Therapeutin, daß er dazu neigt, sich in totales Schweigen zurückzuziehen. Bei Katrin, daß sie dazu neigt, Michael aggressiv nach den Ursachen des Konfliktes zu fragen.

8. Outer Space Physics

**Wenn die Probe reich an Philoben-Gas ist,
ist sie mehr als 22 Grad warm**

I. Causal Forward
Weltraumforscher entdecken im Jahr 4000 einen neuen belebten Planeten in einer fremden Galaxie. Die Forscher widmen sich zuerst der Biophysik des Planeten. An vielen Stellen enthält die Luft das auf der Erde unbekannte Philoben-Gas. Dieses Gas absorbiert Lichtenergie und setzt sie als Wärme frei. Daher sind Stellen, die reich an Philoben-Gas sind, mehr als 22 Grad warm.

II. Causal Backward
Weltraumforscher entdecken im Jahr 4000 einen neuen belebten Planeten in einer fremden Galaxie. Die Forscher widmen sich zuerst der Biophysik des Planeten. An vielen Stellen enthält die Luft das auf der Erde unbekannte Philoben-Gas. Philoben-Gas

entsteht durch eine spontane chemische Reaktion in der Atmosphäre des Planeten, sobald die Temperatur 22 Grad übersteigt.

III. Noncausal
Weltraumforscher entdecken im Jahr 4000 einen neuen belebten Planeten in einer fremden Galaxie. Die Forscher widmen sich zuerst der Biophysik des Planeten. An vielen Stellen enthält die Luft das auf der Erde unbekannte Philoben-Gas.

9. Alarm Equipment

Wenn das Flutlicht an ist,
heult die Sirene

I. Causal Forward
Paul arbeitet als Wachmann in einem Versicherungsgebäude. Seine besondere Aufmerksamkeit gilt dem zentralen Tresorraum im Erdgeschoß. Dort ist eine sehr empfindliche Alarmanlage installiert. Betritt z.B. jemand unangemeldet den Vorraum zum Tresorraum, wird das Flutlicht angeschaltet. Das Flutlicht aktiviert die Sirene der Alarmanlage, die dann zu heulen beginnt.

II. Causal Backward
Paul arbeitet als Wachmann in einem Versicherungsgebäude. Seine besondere Aufmerksamkeit gilt dem zentralen Tresorraum im Erdgeschoß. Dort ist eine sehr empfindliche Alarmanlage installiert. Betritt z.B. jemand unangemeldet den Vorraum zum Tresorraum, beginnt die Sirene der Alarmanlage zu heulen. Das Heulen der Sirene löst die Anschaltung des Flutlichtes aus.

III. Noncausal
Paul arbeitet als Wachmann in einem Versicherungsgebäude. Seine besondere Aufmerksamkeit gilt dem zentralen Tresorraum im Erdgeschoß. Dort ist eine sehr empfindliche Alarmanlage installiert. Betritt z.B. jemand unangemeldet den Vorraum zum Tresorraum, wird das Flutlicht angeschaltet und die Sirene der Alarmanlage beginnt zu heulen.

Appendix B: Correlations in Experiments 1-3

Table B1: Correlations over Items for Causal and Noncausal Conditionals in Experiments 1-3

Correlations over Items											
	MT_d	AC_d	DA_d	MP_i	MT_i	AC_i	DA_i	*P(q|p)*	*P(p→q)*	Excep.	Alt.
MP_d – causal	0.81*	-0.05	0.25	0.68*	0.63*	-0.09	0.16	0.65*	0.65*	-0.72*	-0.07
MP_d – noncausal	0.78*	0.29	0.47*	0.33	0.37	0.01	0.04	0.29	0.38	-0.34	-0.16
MT_d – causal		0.16	0.26	0.58*	0.55*	0.06	0.26	0.55*	0.58*	-0.65*	-0.16
MT_d – noncausal		0.18	0.28	0.39	0.53*	0.10	0.19	0.22	0.17	-0.34	0.00
AC_d – causal			0.63*	0.00	-0.06	0.74*	0.58*	0.19	0.14	-0.10	-0.79*
AC_d – noncausal			0.90*	0.55*	0.36	0.53*	0.38	0.57*	0.53*	-0.55*	-0.75*
DA_d – causal				0.27	0.22	0.58*	0.63*	0.42*	0.36*	-0.14	-0.61*
DA_d – noncausal				0.59*	0.42	0.43	0.34	0.47	0.55*	-0.60*	-0.80*
MP_i – causal					0.75*	-0.09	0.28	0.69*	0.76*	-0.77*	-0.16
MP_i – noncausal					0.62*	0.71*	0.65*	0.50*	0.48*	-0.53*	-0.58*
MT_i – causal						0.02	0.43*	0.61*	0.63*	-0.68*	-0.23
MT_i – noncausal						0.53*	0.62*	0.50*	0.21	-0.67*	-0.27*
AC_i – causal							0.75*	0.11	0.02	0.04	-0.71*
AC_i – noncausal							0.69*	0.51*	0.41	-0.53*	-0.57*
DA_i – causal								0.44*	0.38*	-0.30	-0.75*
DA_i – noncausal								0.36	0.23	-0.49*	-0.46
P(q|p) - causal									0.81*	-0.73*	-0.30
P(q|p) - noncausal									0.85*	-0.54*	-0.46
P(p→q) – causal										-0.76*	-0.26
P(p→q)–noncausal										-0.45	-0.59*
Excep.- causal											0.22
Excep-noncausal											0.71*

Note. Cell entries show correlations over items for causal (n = 36) and noncausal conditionals (n = 18). Correlations are computed over acceptance rates of inferences and estimates of subjective conditional probability, estimates of degree of belief in the conditional, estimates of availability of exceptional situations and of alternative situations.

MP = modus ponens, MT = modus tollens, AC = acceptance of the consequent, DA = denial of the antecedent. The instruction groups are symbolized through a *d* (deduction) or an *i* (induction).

P(q|p) = subjective conditional probability of consequent given antecedent, $P(p \rightarrow q)$ = subjective degree of belief in the conditional, Excep. = availability of exceptional situations, Alt. = availability of alternative situations . * $p < 0.05$.

Appendix C: Instructions and Examples of Items

List of Instructions and Examples of Items:

1. Experiment 1, 2 and 4: Instruction
2. Experiment 1: Example of an Original Problem
3. Experiment 2: Example of an Original Problem
4. Experiment 3 and 5: Instruction for the Deduction Group
5. Experiment 3: Instruction for the Induction Group
6. Experiment 3: Example of an Original Problems (Induction Group)
7. Experiment 3: Example of an Original Problems (Deduction Group)
8. Experiment 4: Example of an Original Problem
9. Experiment 5: Example of an Original Problem

1. Experiment 1, 2 and 4: Instruction

Liebe Teilnehmerin, lieber Teilnehmer!

In diesem kurzen Experiment zeigen wir Ihnen 4 von uns erdichtete kurze Geschichten und stellen Ihnen zu jeder Geschichte einige Fragen. Es gibt keine falschen oder richtigen Antworten, wir interessieren uns für Ihre persönliche Einschätzung!

Die Bearbeitung wird ca. 10 Minuten in Anspruch nehmen.

Mein Alter: [ ]

Mein Geschlecht: ○ weiblich ○ männlich

○ Ich mache zum ersten Mal bei dieser Untersuchung mit.
○ Ich habe bereits einmal an dieser Untersuchung teilgenommen.

WICHTIGER HINWEIS:
Bitte benutzen Sie auf keinen Fall den Zurück-Button Ihres Browsers während des Experimentes!

[Los geht's!]

2. Experiment 1: Example of an Original Problem

Aufgabe 6 von 6

Weltraumforscher entdecken im Jahr 4000 einen neuen belebten Planeten in einer fremden Galaxie. Die Forscher widmen sich zuerst der Biophysik des Planeten. An vielen Stellen enthält die Luft das auf der Erde unbekannte Philoben-Gas. Dieses Gas absorbiert Lichtenergie und setzt sie als Wärme frei. Daher sind Stellen, die reich an Philoben-Gas sind, mehr als 22 Grad warm.

Julia ist Physikerin auf der Erde. Sie nimmt an, daß generell gilt:

Wenn die Probe reich an Philoben-Gas ist,
ist sie mehr als 22 Grad warm.

Frage 1: Aus den Proben der Weltraumforscher wird eine beliebige Probe per Zufall herausgezogen. Es stellt sich heraus, daß diese Probe reich an Philoben-Gas ist.

Für wie wahrscheinlich halten Sie es, daß diese Probe mehr als 22 Grad warm ist?

Bitte geben Sie eine Zahl zwischen **0 (völlig unmöglich)** und **100 (absolut sicher)** an:

Frage 2: Beschreibt Julias Annahme Ihrer Meinung nach eine kausale Beziehung?

○ **Nein**, die Annahme beschreibt keine kausale Beziehung.

Falls **ja**, schätzen Sie bitte die **Stärke** dieser kausalen Beziehung:

sehr schwache kausale Beziehung ○ ○ ○ ○ ○ **sehr starke** kausale Beziehung

---

Frage 3: Gehen Sie für diese Frage bitte davon aus, daß Julias Annahme zutrifft!
Können Sie sich Umstände vorstellen, durch die der folgende Fall möglich ist?

Die Probe ist mehr als 22 Grad warm,
aber sie ist nicht reich an Philoben.

○ **Nein,** das kann ich mir nicht vorstellen. (weiter zu Frage 4)

Ja, das kann ich mir vorstellen. Ich kann mir:

sehr wenige ○ ○ ○ ○ ○ **sehr viele**

unterschiedliche Situationen vorstellen, in denen dieser Fall vorkommen kann.

Bitte beschreiben Sie in wenigen Stichworten eine solche Situation.
Bitte beschränken Sie sich auf eine einzige!

Frage 4: Gehen Sie für diese Frage bitte davon aus, daß Julias Annahme zutrifft!
Können Sie sich Umstände vorstellen, durch die der folgende Fall möglich ist?

Eine Probe ist reich an Philoben,
aber sie ist nicht mehr als 22 Grad warm.

○ **Nein,** das kann ich mir nicht vorstellen. (weiter zur nächsten Aufgabe)

Ja, das kann ich mir vorstellen. Ich kann mir:

sehr wenige ○ ○ ○ ○ ○ **sehr viele**

unterschiedliche Situationen vorstellen, in denen dieser Fall vorkommen kann.

Bitte beschreiben Sie in wenigen Stichworten eine solche Situation.
Bitte beschränken Sie sich auf eine einzige!

Weiter

3. Experiment 2: Example of an Original Problem

Aufgabe 1 von 6

Biologen einer amerikanischen Universität haben vor kurzer Zeit eine neue tropische Pflanze entdeckt und sie Pherotelia getauft. Eine besondere Käferart, die Blaupunktkäfer, bringen die Pherotelia zum Blühen, indem sie ihre Eier in ihre Blätter und Stiele legen.

Stefanie ist eine deutsche Biologin. Sie nimmt an, daß generell gilt:

> **Wenn eine Pherotelia blüht,**
> **findet man auf der Pflanze Blaupunktkäfer.**

Frage 1: Für wie wahrscheinlich halten Sie es, daß Stefanies Annahme zutrifft?

Bitte geben Sie eine Zahl zwischen **0 (völlig unmöglich)** und **100 (absolut sicher)** an: [ ]

Frage 2: Beschreibt Stefanies Annahme Ihrer Meinung nach eine kausale Beziehung?

○ **Nein**, die Annahme beschreibt keine kausale Beziehung.

Falls **ja**, schätzen Sie bitte die **Stärke** dieser kausalen Beziehung:

sehr schwache kausale Beziehung ○ ○ ○ ○ ○ **sehr starke** kausale Beziehung

---

Frage 3: Gehen Sie für diese Frage bitte davon aus, daß Stefanies Annahme zutrifft!
Können Sie sich Umstände vorstellen, durch die der folgende Fall möglich ist?

> **Eine Pherotelia blüht,**
> **aber man findet auf ihr keine Blaupunktkäfer.**

○ **Nein,** das kann ich mir nicht vorstellen. (weiter zu Frage 4)

Ja, das kann ich mir vorstellen. Ich kann mir:

sehr wenige ○ ○ ○ ○ ○ **sehr viele**

unterschiedliche Situationen vorstellen, in denen dieser Fall vorkommen kann.

Bitte beschreiben Sie in wenigen Stichworten eine solche Situation.
Bitte beschränken Sie sich auf eine einzige!

[ ]

Frage 4: Gehen Sie für diese Frage bitte davon aus, daß Stefanies Annahme zutrifft!
Können Sie sich Umstände vorstellen, durch die der folgende Fall möglich ist?

> **Auf einer Pherotelia findet man Blaupunktkäfer,**
> **aber sie blüht nicht.**

○ **Nein,** das kann ich mir nicht vorstellen. (weiter zur nächsten Aufgabe)

Ja, das kann ich mir vorstellen. Ich kann mir:

sehr wenige ○ ○ ○ ○ ○ **sehr viele**

unterschiedliche Situationen vorstellen, in denen dieser Fall vorkommen kann.

Bitte beschreiben Sie in wenigen Stichworten eine solche Situation.
Bitte beschränken Sie sich auf eine einzige!

Weiter

4. Experiment 3 and 5: Instruction for the Deduction Group

Liebe Teilnehmerin, lieber Teilnehmer!

Die Aufgaben dieser Untersuchung haben folgende Form:

Annahme:	"Wenn Anna nach Italien reist, fährt sie mit dem Zug"
Tatsache:	Anna reist nach Italien
Schlußfolgerung:	Anna fährt mit dem Zug

Stellen Sie sich vor, die Aufgaben wären Teil eines Gespräches oder einer Diskussion. Ihr Gegenüber zieht aus der Annahme und der Tatsache die genannte Schlußfolgerung.

Bitte prüfen Sie, ob die Argumentation logisch stimmig ist. Beurteilen Sie **nicht**, ob die Annahme und die Tatsache zutreffen, sondern ob die Schlußfolgerung **logisch zwingend** aus ihnen folgt. Das heißt, wenn die Annahme und die Tatsache wahr sind, kann die Schlußfolgerung nicht falsch sein.

Anschließend geben Sie bitte an, wie sicher Sie sich bei dieser Entscheidung fühlen.

Eine Aufgabe könnte demnach komplett so aussehen:
(Dies ist nur ein Beispiel. Wenn Sie möchten, können sie bereits Antworten anklicken - sie brauchen aber nicht.)

Annahme :	"Wenn Anna nach Italien reist, fährt sie mit dem Zug"
Tatsache:	Anna reist nach Italien
Schlußfolgerung:	Anna fährt mit dem Zug

Frage 1: Ist die Schlußfolgerung logisch gültig?

nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?

unsicher ○ ○ ○ ○ sicher

Bitte beachten Sie, daß die Annahme **nicht** unbedingt **umkehrbar** ist. "Wenn Anna nach Italien reist, fährt sie mit dem Zug" ist nicht gleichbedeutend mit: "Wenn Anna mit dem Zug reist, fährt sie nach

Italien".

Die Untersuchung besteht insgesamt aus 3 Aufgaben, jede Aufgabe ist auf einer Seite dargestellt. Die Bearbeitung wird ca. 10 - 15 Minuten in Anspruch nehmen.

Mein Alter:

Mein Geschlecht: ○ weiblich ○ männlich

○ Ich mache zum ersten Mal bei dieser Untersuchung mit.

○ Ich habe bereits einmal an dieser Untersuchung teilgenommen.

WICHTIGER HINWEIS:
Bitte benutzen Sie auf keinen Fall den Zurück-Button Ihres Browsers während des Experimentes!

Los geht's!

5. Experiment 3: Instruction for the Induction Group

Liebe Teilnehmerin, lieber Teilnehmer!

Die Aufgaben dieser Untersuchung haben folgende Form:

Annahme:	"Wenn Anna nach Italien reist, fährt sie mit dem Zug"
Tatsache:	Anna reist nach Italien
Schlußfolgerung:	Anna fährt mit dem Zug

Stellen Sie sich vor, die Aufgaben wären Teil eines Gespräches oder einer Diskussion. Ihr Gegenüber zieht aus der Annahme und der Tatsache die genannte Schlußfolgerung.

Bitte beurteilen Sie, ob Sie die Argumentation **überzeugend** finden. Würden Sie der Argumentation Ihres Gegenübers folgen und die Schlußfolgerung akzeptieren?

Anschließend geben Sie bitte an, wie sicher Sie sich bei dieser Entscheidung fühlen.

Eine Aufgabe könnte demnach komplett so aussehen:
(Dies ist nur ein Beispiel. Wenn Sie möchten, können sie bereits Antworten anklicken - sie brauchen aber nicht.)

Annahme :	"Wenn Anna nach Italien reist, fährt sie mit dem Zug"
Tatsache:	Anna reist nach Italien
Schlußfolgerung:	Anna fährt mit dem Zug

Frage 1: Finden Sie diese Schlußfolgerung plausibel?

nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?

unsicher ○ ○ ○ ○ sicher

Bitte beachten Sie, daß die Annahme **nicht** unbedingt **umkehrbar** ist. "Wenn Anna nach Italien reist, fährt sie mit dem Zug" ist nicht gleichbedeutend mit: "Wenn Anna mit dem Zug reist, fährt sie nach Italien".

Die Untersuchung besteht insgesamt aus 3 Aufgaben, jede Aufgabe ist auf einer Seite dargestellt. Die Bearbeitung wird ca. 10 - 15 Minuten in Anspruch nehmen.

Mein Alter: [ ]

Mein Geschlecht: ○ weiblich ○ männlich

○ Ich mache zum ersten Mal bei dieser Untersuchung mit.

○ Ich habe bereits einmal an dieser Untersuchung teilgenommen.

WICHTIGER HINWEIS:
Bitte benutzen Sie auf keinen Fall den Zurück-Button Ihres Browsers während des Experimentes!

[Los geht's!]

6. Experiment 3: Example of an Original Problem (Induction Group)

Aufgabe 2 von 3

Weltraumforscher entdecken im Jahr 4000 einen neuen belebten Planeten in einer fremden Galaxie. Die Forscher widmen sich zuerst der Biophysik des Planeten. An vielen Stellen enthält die Luft das auf der Erde unbekannte Philoben-Gas. Dieses Gas absorbiert Lichtenergie und setzt sie als Wärme frei. Daher sind Stellen, die reich an Philoben-Gas sind, mehr als 22 Grad warm.

Julia ist Physikerin auf der Erde. Sie nimmt an, daß generell gilt:

> **Wenn die Probe reich an Philoben-Gas ist,**
> **ist sie mehr als 22 Grad warm.**

Julia erzählt Kollegen von vier Gasproben, die gestern untersucht wurden:

---

1. Probe

Tatsache: Die Probe war reich an Philoben
Julia schließt daraus: Die Probe war mehr als 22 Grad warm

Frage 1: Finden Sie Julias Schlußfolgerung plausibel?

nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?

unsicher ○ ○ ○ ○ sicher

---

2. Probe

Tatsache: Die Probe war mehr als 22 Grad warm
Julia schließt daraus: Die Probe war reich an Philoben

Frage 1: Finden Sie Julias Schlußfolgerung plausibel?

nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?

unsicher ○ ○ ○ ○ sicher

---

3. Probe

Tatsache: Die Probe war weniger als 22 Grad warm
Julia schließt daraus: Die Probe war nicht reich an Philoben

Frage 1: Finden Sie Julias Schlußfolgerung plausibel?

nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?

unsicher ○ ○ ○ ○ sicher

---

4. Probe

Tatsache: Die Probe war nicht reich an Philoben
Julia schließt daraus: Die Probe war weniger als 22 Grad warm

Frage 1: Finden Sie Julias Schlußfolgerung plausibel?

nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?

unsicher ○ ○ ○ ○ sicher

Weiter

7. Experiment 3: Example of an Original Problem (Deduction Group)

Aufgabe 2 von 3

In einer abgelegenen chinesischen Provinz hat die Chemiefirma "Science United" in den letzten Jahren ein neues Pflanzenschutzmittel getestet. Die Forscher, die den Versuch überwachen, haben jetzt festgestellt, daß das Pflanzenschutzmittel bei den einheimischen Kaninchen eine bestimmte neue DNA-Mutation auslöst. Diese Mutation führt dazu, daß im Körper des Kaninchens eine neue Substanz namens Natrolsan gebildet wird.

Angela arbeitet für die Chemiefirma "Biopharma" in der gleichen Region. Sie nimmt an, daß generell gilt:

> **Wenn bei einem Kaninchen Natrolsan nachgewiesen wird, findet man die DNA-Mutation.**

Angela diskutiert mit Kollegen über vier Kaninchen, die in der letzten Woche untersucht worden sind:

---

1. Kaninchen

Tatsache: Es wurde Natrolsan nachgewiesen
Angela schließt daraus: Man fand die DNA-Mutation

Frage 1: Ist Angelas Schlußfolgerung logisch gültig?
nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?
unsicher ○ ○ ○ ○ sicher

---

2. Kaninchen

Tatsache: Man fand die DNA-Mutation
Angela schließt daraus: Es wurde Natrolsan nachgewiesen

Frage 1: Ist Angelas Schlußfolgerung logisch gültig?
nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?
unsicher ○ ○ ○ ○ sicher

---

3. Kaninchen

Tatsache: Man fand die DNA-Mutation nicht
Angela schließt daraus: Es wurde kein Natrolsan nachgewiesen

Frage 1: Ist Angelas Schlußfolgerung logisch gültig?
nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?
unsicher ○ ○ ○ ○ sicher

---

4. Kaninchen

Tatsache: Es wurde kein Natrolsan nachgewiesen
Angela schließt daraus: Man fand die DNA-Mutation nicht

Frage 1: Ist Angelas Schlußfolgerung logisch gültig?
nein ○ ○ ja

Frage 2: Wie sicher sind Sie sich bei Ihrem Urteil?
unsicher ○ ○ ○ ○ sicher

Weiter

8. Experiment 4: Example of an Original Problem

Aufgabe 1 von 4

Weltraumforscher entdecken im Jahr 4000 einen neuen belebten Planeten in einer fremden Galaxie. Die Forscher widmen sich zuerst der Biophysik des Planeten. An vielen Stellen enthält die Luft das auf der Erde unbekannte Philoben-Gas. Philoben-Gas entsteht durch eine spontane chemische Reaktion in der Atmosphäre des Planeten, sobald die Temperatur 22 Grad übersteigt.

Julia ist Physikerin auf der Erde. Sie nimmt an, daß generell gilt:

Wenn die Probe mehr als 22 Grad warm ist,
ist sie reich an Philoben-Gas.

Die Forscher sammeln an möglichst vielen Stellen Gasproben der Luft und erheben Temperatur und Druck der jeweiligen Umgebung für die Analyse Zuhause. Ihre Ergebnisse von 2000 Proben sehen folgendermaßen aus:

90 Proben waren mehr als 22 Grad warm und reich an Philoben.
10 Proben waren mehr als 22 Grad warm und nicht reich an Philoben.
950 Proben waren weniger als 22 Grad warm und reich an Philoben.
950 Proben waren weniger als 22 Grad warm und nicht reich an Philoben.

Frage 1: Für wie wahrscheinlich halten Sie es, daß Julias Annahme zutrifft?

Bitte geben Sie eine Zahl zwischen **0 (völlig unmöglich)** und **100 (absolut sicher)** an:

Frage 2: Beschreibt Julias Annahme Ihrer Meinung nach eine kausale Beziehung?

○ **Nein**, die Annahme beschreibt keine kausale Beziehung.

Falls **ja**, schätzen Sie bitte die **Stärke** dieser kausalen Beziehung:

sehr schwache kausale Beziehung ○ ○ ○ ○ ○ **sehr starke** kausale Beziehung

---

Frage 3: Gehen Sie für diese Frage bitte davon aus, daß Julias Annahme zutrifft!
Können Sie sich Umstände vorstellen, durch die der folgende Fall möglich ist?

Die Probe ist mehr als 22 Grad warm,
aber sie ist nicht reich an Philoben.

○ **Nein,** das kann ich mir nicht vorstellen. (weiter zu Frage 4)

Ja, das kann ich mir vorstellen. Ich kann mir:

sehr wenige ○ ○ ○ ○ ○ **sehr viele**
unterschiedliche Situationen vorstellen, in denen dieser Fall vorkommen kann.

Bitte beschreiben Sie in wenigen Stichworten eine solche Situation.
Bitte beschränken Sie sich auf eine einzige!

Frage 4: Gehen Sie für diese Frage bitte davon aus, daß Julias Annahme zutrifft!
Können Sie sich Umstände vorstellen, durch die der folgende Fall möglich ist?

Eine Probe ist reich an Philoben,
aber sie ist nicht mehr als 22 Grad warm.

○ **Nein,** das kann ich mir nicht vorstellen. (weiter zur nächsten Aufgabe)

Ja, das kann ich mir vorstellen. Ich kann mir:

sehr wenige ○ ○ ○ ○ ○ **sehr viele**

unterschiedliche Situationen vorstellen, in denen dieser Fall vorkommen kann.

Bitte beschreiben Sie in wenigen Stichworten eine solche Situation.
Bitte beschränken Sie sich auf eine einzige!

Weiter

9. Experiment 5: Example of an Original Problem

Aufgabe 4 von 4

In einem großen Labor für Tiermedizin in Australien ist vor kurzem eine neue allergische Krankheit bei Hunden entdeckt worden. Die Forscher haben die Krankheit Midosis getauft. Inzwischen wissen die Wissenschaftler schon viele Details über Midosis. Im Blut der betroffenen Tiere entsteht eine bisher unbekannte Substanz namens Xathylen, diese Substanz führt zu den vielfältigen Symptomen der Midosis.

Beate ist eine niedergelassene Tiermedizinerin. Sie nimmt an, daß generell gilt:

Wenn sich Xathylen im Blut eines Hundes nachweisen läßt,
dann leidet er unter Midosis.

Für 2000 Hunde, die das Labor untersucht hat, gilt folgendes:

900 Hunde hatten Xathylen im Blut und litten unter Midosis.
900 Hunde hatten Xathylen im Blut und litten nicht unter Midosis.
100 Hunde hatten kein Xathylen im Blut und litten unter Midosis.
100 Hunde hatten kein Xathylen im Blut und litten nicht unter Midosis.

Beate diskutiert mit einem Kollegen über vier Fälle, die dieser Kollege untersucht hat:

---

1. Hund
Tatsache: Im Blut des Hundes ließ sich Xathylen nachweisen
Beate schließt daraus: Der Hund litt unter Midosis

Frage 1: Ist Beates Schlußfolgerung logisch gültig?

nein ○ ○ ja

---

2. Hund
Tatsache: Der Hund litt unter Midosis
Beate schließt daraus: Im Blut des Hundes ließ sich Xathylen nachweisen

Frage 1: Ist Beates Schlußfolgerung logisch gültig?

nein ○ ○ ja

---

3. Hund
Tatsache: Der Hund litt nicht unter Midosis
Beate schließt daraus: Im Blut des Hundes ließ sich kein Xathylen nachweisen

Frage 1: Ist Beates Schlußfolgerung logisch gültig?

nein ○ ○ ja

---

4. Hund

Tatsache: Im Blut des Hundes ließ sich kein Xathylen nachweisen

Beate schließt daraus: Der Hund litt nicht unter Midosis

Frage 1: Ist Beates Schlußfolgerung logisch gültig?

nein ○ ○ ja

Weiter