This is a section of doi:10.7551/mitpress/11252.001.0001

The Handbook of Rationality

Edited by: Markus Knauff, Wolfgang Spohn

Citation:

The Handbook of Rationality
Edited by: Markus Knauff, Wolfgang Spohn
DOI: 10.7551/mitpress/11252.001.0001
ISBN (electronic): 9780262366175
Publisher: The MIT Press
Published: 2021

Funding for the open access edition was provided by the MIT Libraries Open Monograph Fund.

The MIT Press

4.6 Probabilities and Conditionals

Klaus Oberauer and Danielle Pessach

Summary

What do conditionals have to do with probability? After a short introduction, the second section of this chapter will explain this connection and how it laid the groundwork for the "probabilistic turn" in the psychology of conditional reasoning. The rest of the chapter is devoted to two questions: (a) What is the *meaning* of conditionals, and (b) how do we *reason* with conditionals? We will answer these two questions from the perspective of probabilistic accounts of conditional reasoning. The third section is concerned with the probabilistic interpretation of the conditional, whereas the fourth section will cover current approaches to reasoning with (probability-)conditionals.

1. Conditionals and Rationality

Conditionals lie at the heart of human rationality. We use them to assert how states and events depend on each other—causally, logically, deontically, or in other ways (e.g., "If you want to learn more about conditionals, you should read this chapter"). We use them to express hypothetical thinking (e.g., "If we reduce CO_2 emissions now, we can still keep global warming below 2°C"). They figure centrally in two major systems for formalizing rationality, formal logic and the probability calculus: first, "if–then"—defined as the material conditional—was included as one of the junctors linking elementary propositions in propositional logic. Subsequently, the interpretation of "if–then" expressions as material conditionals has been criticized, leading to a "probabilistic turn" that links conditionals closely to conditional probabilities. This chapter reviews the arguments motivating this "probabilistic turn" and the related psychological research on how people interpret and reason with conditionals.

2. Why Turn to Probabilities?

The groundwork for the probabilistic turn was laid by the inaptitude of classical logic[1] to account for how people interpret conditionals, as well as for how they reason with them. Classical logic conceptualizes a conditional $A \rightarrow C$ (read as: "If A then C," where A is called the *antecedent* and C is called the *consequent*) as the *material conditional*. Whether a material conditional is true or false is defined by the truth-values of its constituents; it is therefore fully truth-functional. A truth function can be represented as a truth table, which maps the truth-value of a compound sentence such as "If A then C" as a function of the truth conditions of its components A and C. Table 4.6.1 shows the truth table for the material conditional. The material conditional is false when A is true and C is false and true in all other possible cases. This definition has consequences often characterized as paradoxical. From any true C, we can validly infer $A \rightarrow C$: from "Cats are mammals" we can infer "If cats are fish, then cats are mammals." Another paradox: from $\neg A$ ("not-A"; $\neg$ is used to denote a negation), we can validly infer $A \rightarrow C$. For example, "The moon is square" is clearly false; therefore, "If the moon is square, the president is a good person" and "If the moon is square, the president is a bad person" both follow. In the material-conditional account, $A \rightarrow C$ and $A \rightarrow \neg C$ can both be true.

Psychological research has shown that people do not understand conditionals as material conditionals. One possibility to assess how people understand $A \rightarrow C$ is to ask them to indicate for each possible combination of

Table 4.6.1
Truth table for the material conditional and responses in Johnson-Laird and Tagart (1969)

		Conditional ($A \rightarrow C$)	
A	C	MC	Responses
True	True	True	True
True	False	False	False
False	True	True	Irrelevant
False	False	True	Irrelevant

Note: MC = material conditional. Responses are participants' majority responses.

truth-values for A and C (i.e., for each line of the truth table) whether this combination renders the conditional true or false or is irrelevant to the truth of the conditional. This task is referred to as the *truth table task*. The usual pattern is displayed in table 4.6.1: participants tend to judge the false-antecedent cases as "irrelevant" rather than as cases rendering the conditional true (Evans, Handley, Neilens, & Over, 2008; Johnson-Laird & Tagart, 1969; Schroyens, 2010).

People also do not *reason* with conditionals as dictated by classical logic. In the psychology of reasoning, four simple inference forms—taking a conditional major premise and a categorical minor premise—are frequently studied. They are termed "modus ponens" (MP), "modus tollens" (MT), "affirmation of the consequent" (AC), and "denial of the antecedent" (DA). For example, MP in its abstract form is: from "If A then C" and A follows C. The full list of the four inferences is as follows:

MP: $\dfrac{A \rightarrow C \quad A}{C}$ AC: $\dfrac{A \rightarrow C \quad C}{A}$

MT: $\dfrac{A \rightarrow C \quad \neg C}{\neg A}$ DA: $\dfrac{A \rightarrow C \quad \neg A}{\neg C}$

Note: Statements above the horizontal lines are premises, below the line are conclusions.

While MP and MT are valid according to classical logic, AC and DA are not. However, in a conditional inference task (i.e., a task where participants have to judge the four inference forms as valid or not), people usually do not adhere to that norm: although almost everyone endorses MP as valid (~ 95%), the endorsement rate drops significantly for MT (~ 70%), and the so-called *fallacies* (AC and DA) are accepted as valid with rates well over 50% (e.g., Evans, Newstead, & Byrne, 1993; Evans, Over, & Manktelow, 1993; Schroyens & Schaeken, 2003).

The paradoxes of the material implication, its incompatibility with people's judgment (in the truth table task as well as in the conditional inference task), and the fact that everyday reasoning seldom deals with "true" and "false," but rather with uncertainty in different degrees, have led philosophers as well as psychologists to question the material conditional as an explication of the meaning of "if–then" statements. As an alternative, the "probability conditional" has been suggested (Adams, 1975; Bennett, 2003; Edgington, 1995).

3. Probabilistic Account for the Meaning of Conditionals: The Probability Conditional

The probability conditional is defined by "the Equation":

$$P(A \rightarrow C) = P(C \mid A).$$

The Equation defines the probability of the conditional as the conditional probability of the consequent given the antecedent. It was inspired by an account of how people assess conditionals, known as the *Ramsey test*. In probably one of the most famous footnotes in the history of philosophy, Ramsey wrote, "If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q. . . . We can say they are fixing their degrees of belief in q given p" (Ramsey, 1929/1990, p. 247). So, when we assess $A \rightarrow C$, we *suppose A* by adding it (hypothetically) to our beliefs about the world and assess, from that hypothetical belief state, how likely C is. The outcome of the Ramsey test is the conditional probability of C, given A, $P(C \mid A)$.

Before we continue to assess the merits and drawbacks of the probability conditional, some clarifications and specifications are necessary. First, when we speak of probabilities, we mean *subjective* probabilities, sometimes also called *degrees of belief*.

Second, while probabilities of conditionals equal conditional probabilities, they do not inherit all their (mathematical) features. The following two points should make this distinction clear: (1) conditional probabilities are mathematically defined by the Ratio formula: $P(C \mid A) = P(A \,\&\, C)/P(A)$. This does not apply to conditionals: for the purpose of determining $P(A \rightarrow C)$, $P(C \mid A)$ is defined as the outcome of the Ramsey test, that is, it is the result of assessing the probability of C on the supposition of A. Edgington (1995) gives the following example: we can be thinking about what to do next, A or B, considering what consequences those options would have and how likely they are. Obviously, we can do that without having fixed our degree of belief in A—after all, we are still deliberating whether to make A true. Hence, we can determine $P(C \mid A)$ without having a belief about $P(A)$ (Edgington, 1995, p. 266).

(2) Another consequence of the Ratio formula is that $P(C \mid A)$ is not defined when $P(A) = 0$. Stalnaker (1968) offers an extension of the Ramsey test that deals with cases where $P(A) = 0$ (i.e., counterfactual conditionals). If A is incompatible with our knowledge, we cannot simply add it to our stock of knowledge and still be coherent in what we believe. Stalnaker's extended Ramsey test requires that, before adding A, we make the minimally

necessary adjustments to our knowledge to make it compatible with A. However, this extended Ramsey test has been criticized, mostly because it is impossible to determine which adjustments to one's stock of knowledge are "minimal." Bennett (2003), among others, therefore maintains that when $P(A) = 0$, conditionals—with a few exceptions—are not defined and have no probability.

Now that the Equation is disambiguated, we can turn to its main implication: conditionals have no truth conditions.

3.1 A Triviality Result

The triviality result is Lewis's (1976) proof that there is no combination of cases in the truth table for which the joint probability systematically equals $P(C|A)$. Because any proposition conjoining A and C is defined by its truth-value, which in turn is a function of the truth-values of A and C, it follows that there is no proposition that satisfies the Equation (Lewis, 1976). Hence, if the meaning of a conditional such as "If A then C" is defined by the Equation, it cannot be a proposition (Adams, 1975; Bennett, 2003; Douven & Verbrugge, 2010; Edgington, 1995; Lewis, 1976). The consequences are far-reaching: if we accept the Equation, conditionals do not have truth conditions. This seems to be a high price to pay for an explanation of the conditional. For instance, it would not be meaningful to say that a conditional is true or false, and we could no longer characterize mental states such as believing or doubting a conditional as propositional attitudes (i.e., attitudes about a proposition). Why should we pay that price?

3.2 Arguments for the Probability Conditional

Let us first see how the probability conditional fares with regard to the objections against the material conditional. We saw above that the material conditional has some counterintuitive implications when the antecedent is false or when the consequent is true. Considering an antecedent that is certainly false (i.e., $P(A) = 0$), we can follow Bennett and regard the conditional as undefined or follow Stalnaker's extended Ramsey test and regard the conditional as more or less credible depending on how high $P(C|A)$ is after we have adjusted our beliefs to accommodate A. Either version saves us from having to endorse a conditional simply because its antecedent is false. Considering a case where the antecedent A is just very unlikely, but $P(A) > 0$, the Equation ensures that nothing paradoxical follows: $P(C|A)$ can still take on any value.

The same rationale applies to "true" consequents, when they actually are just very likely. It might well be that Peter always goes for a run before he goes to work, and those who know him think it very likely that he goes for a run. However, "If Peter is injured, he goes for a run before he goes to work" is utterly unlikely. This example demonstrates that a high $P(C)$ does not imply a high $P(C|A)$, so that the probability conditional avoids the first paradox of the material conditional in most circumstances. Only when the consequent is a necessary truth, such that $P(C) = 1$, we are also forced to accept the conditional, because the conditional probability must be 1.[2]

Here comes another paradox of the material conditional: a material conditional "If A then C" is false if and only if A and $\neg C$, so from the falsity of "If A then C," we can infer that A and $\neg C$. Most people would agree that "If Peter is in London, then he is in France" is clearly false, regardless of whether Peter is in London or not. Yet the material-conditional account would require Peter to *be* in London for the conditional to be false. The probability conditional does not have that problem. In order to understand that, we have to clarify a consequence of the fact that conditionals have no truth conditions: when we ask, "What is the probability of $A \rightarrow C$?" we do *not* ask, "What is the probability that it is *true*?" A probability conditional can, of course, be true (when antecedent and consequent have obtained), and it can be false (when the antecedent has obtained but the consequent has not). Therefore, it *does* have a truth-value when the antecedent holds. However, in case the antecedent did not obtain (yet), it is neither true nor false; it is simply more or less likely. So one can restate the question how probable we think $A \rightarrow C$ is as how probable we think it is true *given that it has a truth-value*. The conditional has a truth-value when A is true, and in that case, it is true when $A \& C$ is true. Hence, the probability of $A \rightarrow C$, given that it has a truth-value, is $P(A \& C)/P(A)$, and that, of course, matches the conditional probability (Edgington, 1995). Turning back to our London–France example, we see that we have an extreme case: $P(C|A) = 0$. In this case, we can *say* that the conditional is "false" to express that our degree of belief is very low,[3] but in doing so, we do not assign the conditional a truth-value but refer to the conditional probability $P(C|A)$ instead (Jeffrey, 1991). We use "false" in a pleonastic sense, not referring to the value "false" in the truth table but to the zero probability of the conditional (Over & Baratgin, 2017). There are additional philosophical arguments in favor of the probability conditional (for a more detailed account, see Bennett, 2003). We now turn to the experimental evidence.

3.3 Empirical Evidence for the Probability Conditional

As mentioned above, participants in psychological experiments do not judge that false-antecedent cases render the conditional true but judge those cases to be irrelevant for the truth of the conditional (Johnson-Laird & Tagart, 1969). This is in line with the probabilistic account, because the fact that A is false has no bearing on $P(C|A)$. However, this finding can also be explained by an alternative theory that builds on the material conditional for its assumptions on how people represent conditional statements: mental model theory (Johnson-Laird & Byrne, 2002).[4] This theory is currently the main competitor to the probabilistic approach in the psychology of reasoning, and therefore we briefly introduce it next.

Mental model theory (MMT) assumes that the material conditional expresses the fully elaborated core meaning of conditionals. In addition, the theory specifies how conditionals are mentally represented and the psychological processes operating on them when people make judgments about conditionals or engage in conditional reasoning. MMT states that conditionals are represented as a set of mental models. A complete representation of a "basic" conditional (i.e., a conditional for which no contextual or knowledge factors modulate its core meaning) includes one model for each truth table case that makes the material conditional true (i.e., all cases that are possible, assuming the conditional is true):

$$\begin{array}{cc} A & C \\ \neg A & C \\ \neg A & \neg C \end{array}$$

In this notation, each line is a mental model representing one possible state of the world. Note that only "true" rows of the truth table are represented as models, not the false ones (i.e., "$A \& \neg C$" is not represented because it would render the conditional false; Johnson-Laird & Byrne, 2002). MMT also states that usually people only represent the first model explicitly and add the others implicitly, leaving them unspecified:

$$\begin{array}{cc} A & C \\ [\ldots] & \end{array}$$

The three dots in square brackets are an "elliptical" representation of the unspecified cases. The full set of three models can be fleshed out from the initial representation, but that requires additional mental capacity and therefore is often not done.

These assumptions provide an explanation of why people tend to judge false-antecedent cases as "irrelevant" for the truth-value of conditionals in truth table cases: people fail to flesh out the full set of models and therefore cannot fit the $\neg A$-cases into any mental model of the conditional (Johnson-Laird & Byrne, 2002). However, this explanation has a problem: MMT also predicts—correctly—that people judge the conditional as "false" in case of $A \& \neg C$. Yet, this "false" case of the truth table is not explicitly represented as a mental model, and as such, it has the same status as the $\neg A$-cases with respect to the initial, sparse representation in which only the $A \& C$–case is explicitly represented. Inferring that $A \& \neg C$ is incompatible with the conditional requires fleshing out the fully explicit set of three models (Johnson-Laird & Byrne, 2002). There is no mechanism that can explain how, when only the initial "$A \& C$" model is explicitly represented, $A \& \neg C$ can be distinguished from $\neg A \& C$ and $\neg A \& \neg C$. Therefore, in the truth table task, participants should be predicted to judge all three cases as irrelevant or all three as false.

The classic truth table task provides only indirect evidence for the probabilistic account of conditionals. A more direct test of the Equation would be to see whether participants really adhere to it in their judgments of probabilities.

Evans, Handley, and Over (2003) and Oberauer and Wilhelm (2003) examined whether participants' judgments about the probability of a conditional "If A then C" corresponded to the conditional probability $P(C|A)$ or to the probability according to the material-conditional account (i.e., $1 - P(A \& \neg C)$). Participants were tested on abstract conditionals concerning a set of cards that had, for example, colored letters printed on it (e.g., "If there is an A on the card, it is blue"). For each conditional, they were given the relevant frequencies of the four truth table cases (e.g., of the 100 cards, 50 had a blue A, 10 a red A, 20 a blue B, and 20 a red B printed on it). Participants were then asked to estimate the probability of the conditional. Both studies found strong evidence for the probabilistic account and against the material conditional: the probability of the conditional was highly correlated with $P(C|A)$, and a majority of participants was classified as adhering to the Equation.

Over, Hadjichristidis, Evans, Handley, and Sloman (2007) went a step further: instead of giving the relevant frequencies of the truth table cases, they used everyday conditionals and had participants judge the respective probabilities (with the restriction that they had to sum to 100%). Only the conditional probability (computed from the probability ratings for individual events) was correlated with the rated probability of the conditional. However, Evans and colleagues (2003; Evans, Handley, Neilens, & Over, 2007) as well as Oberauer and Wilhelm (2003) also found evidence of a conjunctive interpretation of conditionals: for a substantial subgroup of participants, $P(A \rightarrow C)$ was mostly influenced by $P(A \& C)$. This pattern of probability judgments is referred to as the "conjunctive pattern."

How can we explain the "conjunctive pattern" of probability judgments? Clearly, people do not think

$A \& C$ to be equal to $A \rightarrow C$, as a person equating the two would be utterly incapable of hypothetical thought.[5] It could be argued that the effect is task specific: in the tasks in which this pattern occurred, participants always had to judge how likely a conditional was true for a randomly drawn subset of cards. If there are participants who regard the conditional as true only if at least one of the cards was an "$A \& C$" card, this could explain the conjunctive pattern. Another possibility, suggested by Evans and colleagues (2003), is that this pattern reflects an incomplete Ramsey test. Participants set out to assess the probability of C on the supposition of A (i.e., once A is added to the stock of knowledge). Doing so requires one to compare the probability of resulting C-cases (i.e., $A \& C$) with that of the resulting $\neg C$-cases (i.e., $A \& C$). If participants stopped halfway during this procedure, omitting the computation of $P(A \& \neg C)$, that could explain why they come up with $P(A \& C)$ as the probability of $A \rightarrow C$. Oberauer, Geiger, Fischer, and Weidenfeld (2007) tested these two explanations but rejected both.

There are two other explanations for the conjunctive pattern: Edgington (2003) suggests that the wording of the question is ambiguous. As made clear above, the probability of a conditional is *not* the probability that it is true. Yet this is exactly what was asked in the experiments. So some participants could have ignored the "true" at the end of the sentence and answered with $P(C|A)$, while others took the question literally and hence responded with the only cases where the conditional is true, that is, the "$A \& C$"–cases. As far as we know, this hypothesis has not been tested directly; however, Oberauer and Wilhelm (2003) found no conjunctive pattern when participants were not required to answer with a probability estimate but simply had to decide whether a conditional was true or false. According to Edgington's explanation, this question should have elicited predominantly conjunctive responses, contrary to what was found. This leads us to the last attempt to explain the conjunctive pattern: the effect could be due to difficult arithmetic operations that are necessary to compute the conditional probabilities. There is some evidence favoring this explanation. As mentioned before, the conjunctive pattern disappeared when participants did not need to compute any probability but simply had to *assess* if their probability was high enough for them to accept the conditional as "true" (Oberauer & Wilhelm, 2003). Similarly, Over and colleagues (2007) did not find evidence of a conjunctive interpretation with their everyday conditionals. Although their participants had to give a numerical estimation of the probability of the conditional, they did not have to calculate them on arbitrary numbers provided by the experimenters but could use their own, not necessarily explicit, estimation.[6] And last but not least, evidence supporting the arithmetic-load explanation comes from Fugard, Pfeifer, Mayerhofer, and Kleiter's study (2011). They used a facilitated-response paradigm: participants were asked about probabilities of a six-faced die landing a certain way in an "*x* out of *y*" format (e.g., "2 out of 6"), where participants could just reuse the frequencies given. For example, frequencies could be "Two sides have a yellow circle, two a yellow square, and two a red circle." The estimate of the probability of the conditional "If the die shows a yellow side, it shows a circle" only requires minimal arithmetic ability (namely, adding the two yellow cases). Fugard and colleagues found that (a) overall, approximately 73% of the participants judged the conditional according to the conditional probability, and (b) this percentage increased over the 70 trials each participant had to go through: 55% of the participants shifted to eventually adopt the conditional probability, most of which had adopted a conjunctive interpretation before.

Overall, the experimental evidence for the interpretation of our everyday conditional as the probability conditional is substantial (for further evidence stemming from negated conditionals, see, e.g., Handley, Evans, & Thompson, 2006). Recently, however, the probability conditional has been criticized for ignoring an important aspect of the meaning of conditionals: relevance.

3.4 Inferentialism

"If Peter is wearing a blue shirt today, the sea levels will rise." This conditional strikes most people as odd at least, because the antecedent has no apparent *relevance* for the consequent. Yet, under the material-conditional and probabilistic account, we are forced to accept it respectively as true or highly probable, simply because sea level rise is highly probable. What is wrong with this conditional is the missing semantic link between the antecedent and the consequent (and we therefore refer to it as a "missing-link conditional," in accordance with Douven, 2015). This is the main point of the critique of inferentialism (Douven, 2015; Krzyżanowska, 2015). Inferentialism states that a conditional is only true if there is a sufficiently strong argument from its antecedent, plus relevant background knowledge, to its consequent (Douven, 2015; Krzyżanowska, Wenmackers, & Douven, 2014). It requires an *inferential connection* between antecedent and consequent, which can be a deductive, abductive, or inductive one (Krzyżanowska et al., 2014).

Skovgaard-Olsen, Singmann, and Klauer (2016) tested this requirement experimentally. They found that participants' judgments were predicted almost perfectly by the Equation for regular conditionals, but this relation failed for missing-link conditionals. To be precise, the relation failed for some participants (who, irrespective of the conditional probability, assigned low probability to missing-link conditionals), while others still adhered to the conditional-probability account. This group pattern was confirmed by Skovgaard-Olsen, Kellen, Hahn, and Klauer (2017), who found that only a minority (17%) of participants adhered to the Equation for missing-link conditionals. Another approach to empirically test inferentialism against the probability conditional is offered by "and-to-if" inferences (also referred to as "centering": from "*A* & *C*" follows "If *A* then *C*"). This inference is valid according to both material- and probability-conditional accounts but invalid for an inferentialist. Although people indeed endorse centering (Cruz, Baratgin, Oaksford, & Over, 2015), that seems to be the case only with conditionals where the antecedent is *relevant* for the consequent and not with missing-link conditionals (Skovgaard-Olsen, Kellen, et al., 2017; Skovgaard-Olsen, Singmann, & Klauer, 2017). The main point of disagreement between proponents of the probability conditional and proponents of inferentialism is whether or not the effect of relevance is of a pragmatic or a semantic nature (i.e., whether or not it affects the meaning of the conditional), a question that is still hotly debated (see Douven, Elqayam, Singmann, & van Wijnbergen-Huitink, 2018; Krzyżanowska, Collins, & Hahn, 2017).

4. Probabilistic Accounts of Reasoning with Conditionals

As mentioned in the introductory section, if one accepts the material conditional as the meaning of conditionals, the usual pattern of people's evaluations of conditional inferences implies that people are extraordinarily bad at reasoning: the only inference they are really good at is modus ponens. Mental model theory (Johnson-Laird & Byrne, 2002), which accepts the material conditional as the normative standard, explains these findings by the failure of people to fully flesh out the set of mental models. MP and AC are endorsed if only the initial model ($A \& C$) is represented. If participants manage to add a further model "$\neg A \& C$," they correctly endorse MP and reject AC and DA. To correctly endorse MT, they would have to further add the model "$\neg A \& \neg C$," because it is the only model of the conditional that is also compatible with the minor premise of the MT argument, "not *C*." Hence, if the fully explicit set of three models is fleshed out, participants accept the valid inferences (MP and MT) and reject the fallacies (AC and DA), which would be the normatively correct response by the material-conditional interpretation.

As we have seen in the previous section, there is strong evidence that people interpret conditional statements according to the probability conditional. How do probabilistic theories account for people's pattern of endorsement and rejection of the four simple conditional inferences? One approach is to understand inferences as Bayesian updating of our knowledge in light of the premises (Oaksford & Chater, 2007). This "purely Bayesian" approach interprets the four inferences as inferring the conditional probabilities of the conclusion, given the minor (categorical) premise. For instance, confronted with a DA argument of the form "If *A* then *C*; not *C*; therefore not *A*," reasoners are assumed to draw on their knowledge about *C* and *A* to assess the probability of "not *A*," given "not *C*." Oaksford, Chater, and Larkin (2000; see also Oaksford & Chater, 2001) were the first to propose this interpretation as a formal reasoning model. To model the person's knowledge, they require three parameters to capture the joint subjective probability distribution of *A* and *C* (i.e., the four truth table cases), from which the respective conditional probabilities can be computed: $e = 1 - P(C|A)$, referred to as the "exceptions parameter," $a = P(A)$ and $b = P(C)$. From this joint probability distribution, MP endorsement rate is predicted by $P(C|A)$, MT by $P(\neg A|\neg C)$, AC by $P(A|C)$, and DA by $P(\neg C|\neg A)$.

Oberauer (2006) compared the Oaksford–Chater model with a formalization of MMT and found that the probabilistic account did not fare well in explaining people's pattern of endorsement and rejection of conditional inferences. The problem for the Oaksford–Chater model is that to explain the high MP acceptance rate, it must estimate the exceptions parameter *e* to a very small value. This implies that the model also predicts high MT acceptance, contrary to the data. As a compromise, the model, when fit to data, has been found to underestimate endorsement of MP and overestimate endorsement of MT inferences (Oberauer, 2006; Schroyens & Schaeken, 2003). Oaksford and Chater (2007, 2013) presented a revision of their model in which they introduce a new exceptions parameter e', derived by updating *e* on a single occurrence of a counterexample to the conditional (i.e., an occurrence of an "*A* and $\neg C$" example). Whereas MP endorsement is still estimated by the "old" exceptions parameter *e*, estimation for all other inferences is based on the new parameter. This solves the MP–MT asymmetry problem, because MP and MT are now decoupled.

Oaksford and Chater (2007, 2013) justify the introduction of e' by arguing that the minor premises given in the context of inferences MT, AC, and DA provide information that is equivalent to presenting a counterexample to the conditional. Therefore, the endorsement rates of these inferences should be predicted with a different exceptions parameter. In our view, this justification is contrived: learning that "not *A*" or that "not *C*" does not provide reasons to question "If *A* then *C*." No amount of people who are not hit by a car, or of people who are not injured, should undermine our belief in the conditional "If a person is hit by a car, they will be injured." Therefore, although the revised model of Oaksford and Chater is better able to accommodate the data, it does so at the price of being less principled than the original model.

Another purely Bayesian approach, similar to the original Oaksford–Chater model, is based on Bayesian networks, a probabilistic graphical model that represents the dependencies between variables as a directed acyclic graph (Pearl, 2009). While the two approaches are similar, they are certainly not equivalent, but due to space restrictions, we cannot go into more detail here (for a detailed account, see chapter 4.2 by Hartmann and chapter 7.1 by Pearl in this handbook). One thing both approaches have in common is that they explain conditional reasoning as an inference from the minor premise to the conclusion.

But then, what is the role of the conditional (i.e., the major premise) in the inference task? Or put differently, what is the difference between inferring *C* from the two premises "If *A* then *C*" and *A* (i.e., the full MP argument) and inferring *C* from *A* alone (i.e., a reduced MP argument), as both are interpreted as estimating $P(C|A)$? Within the Bayesian framework, it is not clear how adding the conditional premise could play a role in knowledge updating. Normally, in Bayesian updating, additional premises are conditionalized on, that is, we assess $P(\text{conclusion}\,|\,\text{all premises})$. However, conditionalizing is not possible when the additional premise is itself a conditional probability. Oaksford and Chater (2007) propose that the effect of adding a conditional premise should be reflected in a decreased exceptions parameter $e = 1 - P(C|A)$, that is, the degree of belief in the sufficiency of the antecedent for the consequent increases. Hartmann and Rafiee Rad (2012) propose that instead of only increasing $P(C|A)$, all parameters capturing the joint probability distribution are adjusted so that the Kullback–Leibler distance between the old and the new distributions is minimized.

Klauer, Beller, and Hütter (2010) investigated the effect of the conditional on the inferences by presenting to participants both reduced inferences (i.e., without the major premise) and full inferences (i.e., the standard MP, MT, AC, and DA). The reduced inferences are viewed as an indication of how much the conclusion is endorsed based *only* on background knowledge. For example, when given the information that the streets are wet, most people would conclude that it probably has rained, even if they haven't been told "If it rains, the streets are wet." To model the effect of adding the conditional to the inference, Klauer and colleagues (2010) developed the dual-source model (DSM), a mixture model that implements Oaksford, Chater, and Larkin's (2000) probabilistic model to account for the reduced inferences (the "knowledge" component in the DSM terminology) but adds two components for the full inferences: (a) four form parameters, which capture to what extent the inference forms (i.e., MP, MT, AC, and DA) are seen as logically valid, and (b) a weighting parameter, which indicates how much weight is given to the form-based information relative to the knowledge-based information. Singmann, Klauer, and Beller (2016) compared the original and the revised Oaksford–Chater model (2000, 2007, 2013), the Bayes nets approach with Kullback–Leibler distance minimization (Hartmann & Rafiee Rad, 2012), and the DSM. They found that the DSM explains substantially more variance than the competing models and also confirmed a unique prediction of the DSM. Singmann and colleagues conclude that the effect of adding a conditional goes beyond updating the knowledge base (i.e., the joint probability distribution). Instead, they argue, adding the conditional to the minor premises provides participants with an inference form that is then evaluated in terms of logical soundness. It is important to note that logical soundness does *not* demand one to refer to bivalued logic: MP and MT are also *probabilistically* valid, whereas DA and AC are not (for an introduction to probability logic, see chapter 4.4 by Pfeifer, this handbook).

Further experiments provide support for the DSM by demonstrating that the form parameters can be selectively influenced by experimental variations of the inference form (e.g., to biconditional, "only if" vs. "if–then," or adding premises). Conversely, the form parameters do not change when the inference form is held constant (Klauer et al., 2010; Singmann et al., 2016). One limitation of the DSM is that, as a measurement model, it serves well to measure the effect of inference form but does not provide an explanation of *how* the validity of the inference forms is assessed.

Insight into what causes the effects of inference form could be gained by looking at a different class of theories, known as "dual-process" theories of reasoning. Various dual-process theories exist, even in the small world of conditional reasoning, but we will focus only on the general idea (for a detailed account, see chapter 2.5 by Klauer, this handbook). As the name suggests, they assume two processes, one (usually called Type 1) fast, automatic, and intuitive, the other (usually called Type 2) slower, controlled, and relying strongly on working-memory capacity. The dual-process framework thus naturally explains reasoning from the inference form in terms of the Type 2 process, whereas the knowledge component can be regarded as a Type 1 process. A candidate for a Type 2 process would be, for example, mental model theory, and indeed, Verschueren, Schaeken, and d'Ydewalle (2005) have proposed a dual-process model that integrates probabilistic inference as a Type 1 process and mental models as a Type 2 process. This model fared very well in a formal comparison of several competing models of conditional inferences (Oberauer, 2006).

But, as we have seen above, mental model theory builds on the material conditional, an assumption that is not supported by the empirical data, so we are left with a model that accounts for conditional inferences but is based on an interpretation of the conditional that is untenable. Geiger and Oberauer (2010) have sketched one potential way to keep the explanatory power of MMT for conditional inferences while avoiding the material-conditional interpretation of "if–then" premises. According to this proposal, the meaning of a conditional is not directly captured by a set of mental models that represent its truth conditions—rather, it is incorporated in procedural knowledge. A conditional statement "If *A* then *C*" is understood as an instruction to add an element *C* to every mental model of *A* in working memory. Conditional reasoning starts with building a mental model of the minor (categorical) premise. In case that model includes *A*, the procedure implementing the conditional adds *C* to it. For instance, when thinking through an MP argument, the reasoner first builds a mental model of *A* in working memory, representing the minor premise. The conditional premise instructs a procedure to augment this model, resulting in the model "$A \& C$," from which the conclusion "*C*" follows immediately. Thinking through an MT argument involves more steps: to represent the minor premise, the first model is "$\neg C$." The reasoner now tentatively augments this model by adding "*A*" and by adding "$\neg A$," effectively reasoning from the supposition that *A*, or the supposition that not *A*. The first augmentation results in a model "$A \& \neg C$," which triggers the procedure that incorporates the conditional premise. That procedure adds "*C*" to the model, in conflict with the "$\neg C$" element, and that conflict signals the contradiction between the supposition of *A* and the premises, which leads to elimination of the "$A \& \neg C$" model. The remaining "$\neg A \& \neg C$" model licenses the conclusion "not *A*."

5. Conclusion

Just rethink for a moment: if you had not known, would you have guessed that the probability of the conditional, judged by laypeople, is approximately equal to the conditional probability? On the one hand, this seems implausible, as calculating conditional probabilities is difficult. On the other hand, uncertainty pervades our knowledge of the world, and we have a rough sense of how likely certain events are to happen and how credible certain claims are. Probabilities are merely a way of expressing our degrees of certainty numerically. Moreover, a hallmark of human rationality is the consideration of hypothetical scenarios (Evans & Over, 2004): when we plan a course of action, we think about what is likely to happen if we do this or that. When we reflect on something that went wrong, we consider whether things would have turned out better if one or another unfortunate event had not happened.

Thinking about hypothetical scenarios does not require the ability to calculate conditional probabilities. It requires to form a representation of the hypothetical state—the supposition—and to use relevant knowledge to infer what is likely to happen, perhaps through a mental simulation (Hegarty, 2004). When people are asked how certain they are of their beliefs about what would happen, they can express their degree of certainty on a coarse scale of subjective probabilities; Khemlani, Lotstein, and Johnson-Laird (2015) have proposed a simple algorithm for arriving at such coarse probability estimates on the basis of mental models. Thinking about what happens, or would have happened, in hypothetical scenarios is an assessment of conditional probabilities: how probable is outcome *C*, given the hypothetical event or action *A*? Speaking about such hypothetical scenarios involves using conditionals (e.g., "If I go home early, my boss won't be pleased. I better stay another hour"). Therefore, a close link between subjective conditional probabilities and the meaning of a conditional statement appears plausible.

The probability conditional and the Equation have received compelling support from studying how people *understand* conditionals (Handley et al., 2006; Oberauer & Wilhelm, 2003; Over et al., 2007). At the same time, the question of how people *reason* with conditionals is still unresolved, both on the normative and the descriptive level. On the normative side, the purely Bayesian approaches have no rational procedure for updating knowledge, given conditional premises. Probability logic could be a possible answer to that problem, but that would mean abandoning the purely Bayesian approach to reasoning. On the descriptive side, people's rates of endorsement of conditional inferences under various conditions are explained only poorly by Bayesian updating models. Dual-source or dual-process models, in which Bayesian updating represents the knowledge-based process, appear more promising. These models require more work devoted to spelling out the form-based process in which reasoners use conditional premises to inform their inferences.

Notes

1. By "classical logic," we are referring to the standard bivalued logic, like, for example, first-order or propositional logic.

2. This has been the source of some criticism of the probability conditional, as we will see later on.

3. Obviously, the same rationale applies to "true" when our degree of belief is very high.

4. There have been various attempts at a substantial revision of mental model theory; see, for example, Barrouillet, Gauffroy, and Lecas (2008) and Johnson-Laird, Khemlani, and Goodwin (2015). For a discussion of the plausibility of these revisions, see Oberauer and Oaksford (2008) and Baratgin et al. (2015), respectively.

5. This person would, after hearing the warning "If you touch the wire, you will get an electric shock," answer "No, that is not true" because he does not touch the wire and is not shocked.

6. We do *not* suggest that people ask themselves first "What is the probability of $A \& C$?" and "What is the probability of $A \& \neg C$?" and then compute a conditional probability (see our remarks on the Ratio formula above) but that people can directly assess $P(C \mid A)$ if they have the relevant background knowledge, as is the case in Over and colleagues' experiment, which used everyday causal conditionals.

References

Adams, E. W. (1975). *The logic of conditionals: An application of probability to deductive logic*. Dordrecht, Netherlands: Reidel.

Baratgin, J., Douven, I., Evans, J. St.B.T., Oaksford, M., Over, D., & Politzer, G. (2015). The new paradigm and mental models. *Trends in Cognitive Sciences*, *19*(10), 547–548.

Barrouillet, P., Gauffroy, C., & Lecas, J. F. (2008). Mental models and the suppositional account of conditionals. *Psychological Review*.

Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford, England: Clarendon Press.

Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in Psychology*, *6*, 192.

Douven, I. (2015). How to account for the oddness of missing-link conditionals. *Synthese*, *194*, 1541–1554.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2018). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology*, *101*, 50–81.

Douven, I., & Verbrugge, S. (2010). The Adams family. *Cognition*, *117*(3), 302–318.

Edgington, D. (1995). On conditionals. *Mind*, *104*(414), 235–329.

Edgington, D. (2003). What if? Questions about conditionals. *Mind & Language*, *18*(4), 380–401.

Evans, J. St. B. T., Handley, S. J., Neilens, H., & Over, D. E. (2007). Thinking about conditionals: A study of individual differences. *Memory & Cognition*, *35*(7), 1772–1784.

Evans, J. St. B. T., Handley, S. J., Neilens, H., & Over, D. (2008). Understanding causal conditionals: A study of individual differences. *Quarterly Journal of Experimental Psychology*, *61*(9), 1291–1297.

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*(2), 321–335.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. (1993). *Human reasoning: The psychology of deduction*. Hillsdale, NJ: Erlbaum.

Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford, England: Oxford University Press.

Evans, J. St. B. T., Over, D. E., & Manktelow, K. I. (1993). Reasoning, decision making and rationality. *Cognition*, *49*(1–2), 165–187.

Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts toward the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 635–648.

Geiger, S. M., & Oberauer, K. (2010). Towards a reconciliation of mental model theory and probabilistic theories of conditionals. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 289–307). Oxford, England: Oxford University Press.

Handley, S. J., Evans, J. St. B. T., & Thompson, V. A. (2006). The negated conditional: A litmus test for the suppositional conditional? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 559–569.

Hartmann, S., & Rafiee Rad, S. (2012). *Updating on conditionals = Kullback–Leibler + causal structure*. Unpublished presentation. Retrieved from https://homepages.uni-regensburg.de/~pfn23853/rffcw/h.pdf

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, *8*(6), 280–285.

Jeffrey, R. (1991). Matter-of-fact conditionals. *Proceedings of the Aristotelian Society, Supplementary Volumes*, *65*, 161–183.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*(4), 646–678.

Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*.

Johnson-Laird, P. N., & Tagart, J. (1969). How implication is understood. *American Journal of Psychology*, *82*(3), 367–373.

Khemlani, S. S., Lotstein, M., & Johnson-Laird, P. N. (2015). Naive probability: Model-based estimates of unique events. *Cognitive Science*, *39*(6), 1216–1258.

Klauer, K. C., Beller, S., & Hütter, M. (2010). Conditional reasoning in context: A dual-source model of probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 298–323.

Krzyżanowska, K. (2015). *Between "if" and "then": Towards an empirically informed philosophy of conditionals* (Unpublished doctoral dissertation). University of Groningen, Netherlands. Retrieved from http://www.rug.nl/research/portal/files/15864231/Title_and_contents_.pdf

Krzyżanowska, K., Collins, P. J., & Hahn, U. (2017). Between a conditional's antecedent and its consequent: Discourse coherence vs. probabilistic relevance. *Cognition*, *164*(Suppl C), 199–205.

Krzyżanowska, K., Wenmackers, S., & Douven, I. (2014). Rethinking Gibbard's riverboat argument. *Studia Logica*, *102*(4), 771–792.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. In W. L. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs* (pp. 129–147). Dordrecht, Netherlands: Springer.

Oaksford, M., & Chater, N. (2001). The probabilistic approach to human reasoning. *Trends in Cognitive Sciences*, *5*(8), 349–357.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, England: Oxford University Press.

Oaksford, M., & Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Thinking & Reasoning*, *19*(3), 346–379.

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(4), 883–899.

Oberauer, K. (2006). Reasoning with conditionals: A test of formal models of four theories. *Cognitive Psychology*, *53*(3), 238–283.

Oberauer, K., Geiger, S. M., Fischer, K., & Weidenfeld, A. (2007). Two meanings of "if"? Individual differences in the interpretation of conditionals. *Quarterly Journal of Experimental Psychology*, *60*(6), 790–819.

Oberauer, K., & Oaksford, M. (2008). What must a psychological theory of reasoning explain? Comment on Barrouillet, Gauffroy, and Lecas (2008). *Psychological Review, 115*(3), 773–778.

Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(4), 680–693.

Over, D. E., & Baratgin, J. (2017). The "defective" truth table: Its past, present, and future. In N. Galbraith, E. Lucas, & D. E. Over (Eds.), *The thinking mind: A Festschrift for Ken Manktelow* (pp. 15–28). London, England: Routledge.

Over, D. E., Hadjichristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, *54*(1), 62–97.

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge, England: Cambridge University Press. (Original work published 2000)

Ramsey, F. P. (1990). *Foundations of mathematics and other logical essays* (R. B. Braithwaite, Ed.). London, England: Routledge. (Original work published 1929)

Schroyens, W. (2010). A meta-analytic review of thinking about what is true, possible, and irrelevant in reasoning from or reasoning about conditional propositions. *European Journal of Cognitive Psychology*, *22*(6), 897–921.

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*(1), 140–149.

Singmann, H., Klauer, K. C., & Beller, S. (2016). Probabilistic conditional reasoning: Disentangling form and content with the dual-source model. *Cognitive Psychology*, *88*, 61–87.

Skovgaard-Olsen, N., Kellen, D., Hahn, U., & Klauer, K. C. (2017). Conditionals, individual variation, and the scorekeeping task. In M. Gunzelmann, A. Howes, T. Tenbrink., & E. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 1084–1089). London, England: Cognitive Science Society. Retrieved from https://philpapers.org/rec/SKOCIV

Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016). The relevance effect and conditionals. *Cognition*, *150*, 26–36.

Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2017). Relevance and reason relations. *Cognitive Science*, *41*, 1202–1215.

Stalnaker, R. C. (1968). A theory of conditionals. In W. L. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs* (pp. 41–55). Dordrecht, Netherlands: Springer.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005). Everyday conditional reasoning: A working memory–dependent tradeoff between counterexample and likelihood use. *Memory & Cognition*, *33*(1), 107–119.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Knauff, Markus, editor. | Spohn, Wolfgang, editor.
Title: The handbook of rationality / edited by Markus Knauff and Wolfgang Spohn.
Description: Cambridge : The MIT Press, 2021. | Includes bibliographical references and index.
Identifiers: LCCN 2020048455 | ISBN 9780262045070 (hardcover)
Subjects: LCSH: Reasoning (Psychology) | Reason. | Cognitive psychology. | Logic. | Philosophy of mind.
Classification: LCC BF442 .H36 2021 | DDC 153.4/3—dc23
LC record available at https://lccn.loc.gov/2020048455