# 6.2 The Suppositional Theory of Conditionals and Rationality

**David E. Over and Nicole Cruz**

**Summary**

In a suppositional theory, a conditional, *if A then B*, is interpreted as *B, supposing A*, and its probability, *P*(*if A then B*), is the conditional probability of *B* given *A*, *P*(*B* | *A*). Conditionals are ubiquitous in reasoning and essential to it. There is strong experimental support for the hypothesis that *P*(*if A then B*) = *P*(*B* | *A*) in people's judgments, and this finding opens up the possibility of giving a unified Bayesian account of human rationality, reasoning, and decision making. Probability theory can be used to define validity in reasoning, and coherence intervals can be specified for the conclusions of inferences. It is rational, by Bayesian standards, to conform to these intervals, and by this measure, people are often rational but sometimes commit fallacies and have biases. Investigating the coherence and incoherence in people's reasoning will give us a deeper understanding of its rationality.

## 1. Conditionals and Supposing

The suppositional theory of conditionals (Edgington, 2014) is a contribution to the Bayesian analysis of human rationality and its understanding of subjective probability judgments as expressing degrees of belief. To state the theory informally, it treats a natural-language indicative conditional, *if A then B*, as equivalent to *B, supposing A*, and it identifies the subjective probability of the conditional, *P*(*if A then B*), with the probability of *B* given *A*, so that *P*(*if A then B*) = *P*(*B* | *A*). This identity, which originates in foundational work on subjective probability theory (de Finetti, 1936/1995, 1937/1964; Ramsey, 1926/1990, 1929/1990), has such deep implications for a Bayesian account of rationality and reasoning that it has been called *the Equation* (Edgington, 1995).

Given the Equation for the natural-language conditional, the normative standard for rational reasoning with it is Bayesian probability theory (chapter 4.5 by Chater & Oaksford, this handbook; Oaksford & Chater, 2007; Over, 2020; Over & Cruz, 2018). In the psychology of reasoning, the Equation becomes *the conditional probability hypothesis*: that people's probability judgments will conform to *P*(*if A then B*) = *P*(*B* | *A*). A suppositional theory of the conditional, as we will interpret it, has both normative and descriptive aspects. It endorses the Equation as setting the correct normative standard for people's conditional reasoning and implies the descriptive conditional probability hypothesis. Supporters of the theory aim to confirm the hypothesis in experimental studies and to investigate how far people comply with the standards of probability theory in their conditional reasoning. Suppositional theory does not imply this compliance will be perfect. Even if people judge that *P*(*if A then B*) = *P*(*B* | *A*), cognitive limitations will sometimes result in biases and fallacious inferences. However, suppositional theory does propose that people's conditional reasoning will be more accurately described by probability theory than by binary extensional logic (Elqayam & Over, 2013).

Reasoning from *A* to *B* (supplemented with background beliefs) can be "summed up" with a conditional assertion *if A then B*, and assertions of *if A then B* can be supported by reasoning from *A* to *B* (plus background information). Thus, any research program implying that *P*(*if A then B*) = *P*(*B* | *A*), as a normative principle and as broadly descriptively adequate, places conditional probability at the center of its account of human reasoning. This identity opens up the possibility of integrating the study of reasoning with that of judgment and decision making, which depends so much on conditional probability, so producing an overall account, normative and descriptive, of human rationality (for developments along these lines, see Oaksford & Chater, 2007, 2020; for the debate in psychology about human rationality, see chapter 1.2 by Evans, this handbook).

## 2. The Conditional Probability Hypothesis

Psychological research has highly corroborated the conditional probability hypothesis as descriptive of people's probability judgments across a wide range of conditionals (Evans, Handley, Neilens, & Over, 2007; Evans, Handley, & Over, 2003; Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011; Oberauer & Wilhelm, 2003; chapter 4.6 by Oberauer & Pessach, this handbook; Over, Hadjichristidis, Evans, Handley, & Sloman, 2007; Pfeifer & Kleiter, 2005; Singmann, Klauer, & Over, 2014). Its intuitive appeal can be appreciated by considering these examples:

(1) If the fair coin is tossed (*F*), then the result will be Heads (*H*).

(2) If the double-headed coin is tossed (*D*), then the result will be Tails (*T*).

The above are *singular conditionals*. They are about specific events: the next toss of the fair, or the double-headed, coin. A *general* conditional is about a whole domain of objects, as in "If an animal is a bird, then it can fly." There are clearly counterexamples to this conditional in its domain, there being birds that cannot fly. A singular conditional with a similar topic would be, "If the next animal we see is a bird, then it will be able to fly." This conditional could be highly probable, depending on where we are in the world, and could turn out to be true. We will focus on singular conditionals like (1) and (2) in this chapter (see Cruz & Oberauer, 2014, on general conditionals).

Intuitively, the probability of (1), *P*(*if F then H*), is .5, and the probability of (2), *P*(*if D then T*), is 0. These are the correct judgments by the Equation, and those that we would predict using the conditional probability hypothesis, since $P(H|F) = .5$ and $P(T|D) = 0$. These judgments are not, however, correct and predictable if the natural-language conditional is interpreted as the *material conditional* of binary extensional logic, *P*(*if A then B*) = $P(A \supset B)$ = *P*(*not-A or B*). The material conditional $A \supset B$ is logically equivalent to the disjunction *not-A or B*, and similarly, *not-A*$\supset$*B* is logically equivalent (*not-not-A* becoming *A*) to *A or B*. However, the identification of *P*(*if A then B*) with *P*(*not-A or B*) is most counterintuitive, and the same can be said for identifying *P*(*if not-A then B*) with *P*(*A or B*). Assume the coins referred to in (1) and (2) are valuable ancient Roman examples in a locked museum display case. Then the probabilities that they will be tossed, *P*(*F*) and *P*(*D*), are very low, making *P*(*not-F*) and *P*(*not-D*), and so *P*(*not-F or H*) and *P*(*not-D or T*), very high. But it is absurd to claim that the probabilities of (1) and (2) are very high merely because these coins are highly unlikely to be tossed, and it is difficult to see how rational reasoning and rational decision making, with natural-language conditionals *if A then B*, can be given a unified treatment when *P*(*if A then B*) = $P(A \supset B)$ = *P*(*not-A or B*).

The logical problem with identifying, for instance, *P*(*if F then H*) and *P*(*not-F or H*) is that it is valid to infer *not-F or H* from *not-F*, which entails that *P*(*not-F*) ≤ *P*(*not-F or H*), forcing *P*(*not-F or H*) to be high whenever *P*(*not-F*) is high. A truth table can be used in classical logic to prove that it is logically valid to infer *A or B* from *A*, an inference termed *or-introduction*. Table 6.2.1 is the truth table for *A or B*, and it clearly shows that *A or B* is true when *A* is true, making the inference from *A* to *A or B* valid, with the further normative result that *P*(*A*) ≤ *P*(*A or B*). Table 6.2.2 is the truth table for the material conditional $A \supset B$ or, equivalently, *not-A or B*, which validly follows from *not-A*. But to claim that it is valid to infer a natural-language conditional *if A then B* from *not-A* is one of the "paradoxes" of the material conditional (Edgington, 2014; Evans & Over, 2004). Its validity entails that *P*(*not-A*) ≤ *P*(*if A then B*), which produces the highly counterintuitive results we have seen in the example above, and people have been found to reject the validity of this inference in their probability judgments (Cruz, Over, & Oaksford, 2017; Pfeifer & Kleiter, 2011; Politzer & Baratgin, 2016).

The only major theory in the psychology of reasoning that closely related (until recently) people's representation of the natural-language conditional to the material conditional was *mental model theory*, with the result that the paradoxes of the material conditional were "valid" in this theory (Johnson-Laird & Byrne,

**Table 6.2.1**
The truth table for the disjunction *A or B*

| *A* | *B* | *A or B* |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 0 | 1 |
| 0 | 1 | 1 |
| 0 | 0 | 0 |

*Note:* 1 = true; 0 = false.

**Table 6.2.2**
The truth table for the material conditional $A \supset B$, equivalent to *not-A or B*

| *A* | *B* | $A \supset B$ |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 0 | 0 |
| 0 | 1 | 1 |
| 0 | 0 | 1 |

*Note:* 1 = true; 0 = false.

1991, pp. 73–74). But counterintuitive implications like those described above have led to a radical revision of mental model theory, in which the paradoxes are no longer considered valid (Johnson-Laird, Khemlani, & Goodwin, 2015). The revision is in a state of development (chapter 2.3 by Johnson-Laird, this handbook), and we will not consider it here (but see Baratgin et al., 2015; Oaksford, Over, & Cruz, 2019; Over, 2020; Over & Cruz, 2018).

Support for the conditional probability hypothesis also disconfirms the claim (Johnson-Laird & Byrne, 1991, p. 74) that a natural-language conditional can be given the same truth table as the material conditional (table 6.2.2). This table entails that $P(A \supset B) = P(\textit{not-A or } B)$, but $P(B \mid A) = P(\textit{not-A or } B)$ only in certain extreme cases (Gilio & Over, 2012), for example, when $P(A) = P(B) = 1$. Four possibilities are displayed in table 6.2.2: *A & B*, *A & not-B*, *not-A & B*, and *not-A & not-B*. As an example, assume we judge that these possible states have the same probability: $P(A \,\&\, B) = P(A \,\&\, \textit{not-}B) = P(\textit{not-}A \,\&\, B) = P(\textit{not-}A \,\&\, \textit{not-}B) = .25$. We now clearly see that $P(B \mid A) = P(A \,\&\, B)/P(A) = .5$, but $P(\textit{not-A or } B) = P(A \,\&\, B) + P(\textit{not-}A \,\&\, B) + P(\textit{not-}A \,\&\, \textit{not-}B) = .75$. More informally, the problem is that *not-A or B* makes a disjunctive statement about the *not-A*-possibilities, with *not-A or B* true in those possibilities, and so the probability of these *not-A*-cases contributes to $P(\textit{not-A or } B)$ but not to $P(B \mid A)$. Lewis (1976) proved a much more general result. He considered conditionals like those in Lewis (1973) and Stalnaker (1968). We might call these Lewis-style conditionals and symbolize them with *if A* $\textit{then}_L$ *B*, which means that *B* holds in the closest possible world (or worlds) in which *A* holds. In this way, *if A* $\textit{then}_L$ *B* can hold in a *not-A*-world, and that implies that $P(\textit{if } A \textit{ then}_L\, B)$ can be affected by $P(\textit{not-}A)$. Because of that, Lewis was able to prove that $P(\textit{if } A \textit{ then}_L\, B)$ cannot generally be identified with $P(B \mid A)$, which depends only on *A*-worlds.

## 3. The Ramsey Test and the de Finetti and Jeffrey Tables

A suppositional conditional is not a material conditional or a Lewis-style conditional (Edgington, 2014). In suppositional theory, *if A then B* means *B, supposing A*, which does not make any statement about, or depend upon, *not-A*-possibilities. We can make the identification that $P(\textit{if } A \textit{ then } B) = P(B \mid A)$ in a suppositional theory, and so this conditional does satisfy the Equation. It can simply be called a *suppositional conditional*, but it has also been termed a probability conditional (Adams, 1998) and a conditional event (de Finetti, 1936/1995, 1937/1964). We can assess $P(\textit{if } A \textit{ then } B)$ for a suppositional conditional by using the *Ramsey test*, which was first described by Ramsey (1929/1990) and later extended by Stalnaker (1968). In this psychological process, people make a judgment about $P(\textit{if } A \textit{ then } B)$ by supposing *A*, while making minimal changes to preserve (or restore) consistency in their beliefs, and then judging their degree of belief in *B* under this supposition, yielding their subjective conditional probability for *B* given *A*, $P(B \mid A)$. The Ramsey test has had a major impact in philosophy, psychology, and other parts of cognitive science (Edgington, 1995; Evans & Over, 2004; Oaksford & Chater, 2007; Pearl, 2013).

We can explain the Ramsey test more fully with an example: assume we are in a car and trying to make a rational decision about which route to take home one particular afternoon. We are discussing the following two conditionals relevant to this decision:

(3) If we take the short city route (*S*), it will be a quick journey (*Q*).

(4) If we take the long countryside route (*L*), it will be an enjoyable journey (*E*).

The short city route home is an uninteresting drive, but the long countryside route can be a pleasant drive. Of course, it is not certain that the city route will be quick and the countryside route enjoyable. A number of uncertain factors can affect these outcomes: traffic jams, changing weather conditions, and accidents. For a rational decision, we must assess both $P(\textit{if } S \textit{ then } Q)$ and $P(\textit{if } L \textit{ then } E)$, and we can do that via the Ramsey test and the Equation. We might make the first Ramsey test on (3), by supposing *S* and making a judgment about how likely *Q* is to follow, and the second on (4), by supposing *L* and making a judgment about how likely *E* is to follow. Perhaps in both tests, we judge that traffic jams and so on are unlikely on this particular day, and so we end the tests by judging that $P(Q \mid S)$ and $P(E \mid L)$ are both high, giving us high confidence in (3) and (4). A rational decision would then depend on our relative desires, that particular afternoon, for a quick drive as opposed to a pleasant one.

Let us say that, after some pondering, we decide that afternoon to take the longer but more pleasant drive, and we set off on the long countryside route. Now our probability judgments change, and we become certain that *S* is false: $P(\textit{not-}S)$ becomes 1. In this case, Ramsey (1929/1990) said that we would lose interest in (3), which he described as becoming "void," and de Finetti (1936/1995, 1937/1964) also took this view and used "void" in this way. We would not use (3) when we were on the countryside route and still less an indicative conditional beginning, "If we are on the city route." We could, though, take an interest in and assert a counterfactual:

(5) If we had taken the short city route home, it would have been a quick journey.

Counterfactuals are a topic of considerable interest in cognitive, social, and developmental psychology. For instance, Kahneman and Miller (1986) established the interesting result that, when the city route is our usual way home, we would feel more regret when we had an accident on the countryside route than when we had an accident on the city route. Our relatively intense regret after an accident on the countryside route and our high confidence in counterfactual (5), which would come from a Ramsey test on (5) similar to the one on (3), would be closely linked. This regret could affect, rationally or irrationally, our future decision making about which route to take. With space limitations, we will, however, continue to focus on indicative conditionals, like (1)–(4), in the rest of this chapter (but see Over, 2017; Over & Cruz, 2019).

In de Finetti's account of indicative conditionals, *if A then B* is true when *A* and *B* hold, false when *A* holds and *B* does not, and "void" when *A* does not hold. These three values can be represented in what has been called a *de Finetti table*, and we display this as table 6.2.3. In traditional research on truth tables in the psychology of reasoning, participants produce a so-called defective table that arguably matches table 6.2.3 and confirms de Finetti's analysis (Baratgin, Over, & Politzer, 2013; Over & Baratgin, 2017; Politzer, Over, & Baratgin, 2010). As we have already noted, Ramsey also used "void" for an indicative conditional with an antecedent known to be false. He added that we lose interest in these indicative conditionals "except as a question about what follows from certain laws or hypotheses" (Ramsey, 1929/1990, p. 155). This suggests that conditionals that follow from logical laws, such as *if A & B then B*, are not void but rather certainly true (Baratgin et al., 2013), and the Ramsey test itself implies that, in general, a conditional with a false antecedent is only void in not having a *factual* truth value, and that it has its conditional probability as a subjective value. A table in which *if A then B* has the value $P(B|A)$ when *A* is false has been called a *Jeffrey table* (after Jeffrey, 1991), and it is found, explicitly or implicitly, in contemporary followers of de Finetti and Ramsey (Cruz & Oberauer, 2014; Over & Baratgin, 2017; Over & Cruz, 2018; Pfeifer & Kleiter, 2009). Table 6.2.4 is the Jeffrey table for *if A then B*.

**Table 6.2.3**
The de Finetti table for *if A then B*

| *A* | *B* | *if A then B* |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 0 | 0 |
| 0 | 1 | V |
| 0 | 0 | V |

*Note:* 1 = true; 0 = false; V = void.

**Table 6.2.4**
The Jeffrey table for *if A then B*

| *A* | *B* | *if A then B* |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 0 | 0 |
| 0 | 1 | $P(B\|A)$ |
| 0 | 0 | $P(B\|A)$ |

*Note:* 1 = true, 0 = false; $P(B|A)$ = the subjective conditional probability of *B* given *A*.

## 4. Rational Reasoning from Degrees of Belief

The suppositional theory of conditionals employs the Ramsey test and de Finetti and Jeffrey tables to give a probabilistic account of conditionals. It rests on the foundational work of Ramsey and de Finetti on conditionals and, more generally, on their identification of people's probability judgments with their degrees of belief. With its focus on degrees of belief, suppositional theory is a prime example of the new Bayesian paradigm in the psychology of reasoning (Elqayam & Over, 2013). Contributions to the traditional paradigm were primarily concerned with inferences from premises that were supposed to be *assumed* true, which meant that these were, in effect, taken to be certain, and the relevant normative standard was binary extensional logic (Johnson-Laird & Byrne, 1991; Over & Cruz, 2018). However, most human inferences, in ordinary affairs and science, are from uncertain beliefs or from hypotheses that, though plausible, are not certain. This uncertainty has to be taken into account in reasoning to arrive at a rational degree of confidence in a conclusion.

For instance, after we decide to take the countryside route home and set off on it, we can use the valid inference form of modus ponens (MP) to infer that we will have an enjoyable journey from (4) as an additional premise. More formally, we would use MP to infer *E* from *if L then E* as the major premise and *L* as the minor premise. We would not, though, treat (4) as certain or simply assume it. Possible *disabling* conditions, like a sudden change in the weather, could make (4) false. By the Ramsey test, (4) has a probability that is equal to $P(E \mid L)$, which is high but is not 1, with 1 representing certainty.

But let us consider first the simpler case of a one-premise inference, rather than a two-premise inference, like MP. In their classic article on conjunction, Tversky and Kahneman (1983) pointed out that, because it is logically valid to infer $B$ from $A \& B$, people should judge that $P(B) \geq P(A \& B)$. Tversky and Kahneman relied on a necessary relation between logical validity and *coherent* probability judgments, which conform to the rules of probability theory. To be incoherent by violating the principle that $P(B) \geq P(A \& B)$ in one's probability judgments is to commit what Tversky and Kahneman called the *conjunction fallacy*.

The coherence of probability judgments can be used to define validity. We can say that a one-premise inference is probabilistically valid, *p-valid*, if and only if the probability of the conclusion can never be coherently lower than the probability of the premise. For an inference with two or more premises, we first define the uncertainty of a statement as 1 minus its probability. For example, when $P(A) = .6$, the uncertainty of $A$ is .4. Then an inference is *p-valid* if and only if, for all coherent probability assignments, the uncertainty of the conclusion is never greater than the sum of the uncertainties of the premises (Adams, 1998). A p-valid inference never increases the uncertainty present in its premises, and it differs from a strong inductive argument in precisely this respect (Evans & Over, 2013).

The one-premise inference from $A \& B$ to $B$ is known as *&-elimination* in logic, and we can ask what rational degree of confidence we should have in $B$ when there is some uncertainty about $A \& B$. The answer given above is that our degree of confidence should be such that $P(B) \geq P(A \& B)$. More formally, for $P(A \& B) = x$, the *coherence interval* for this inference is $[x, 1]$, meaning that $P(B)$ should lie between $x$ and 1 inclusive.

Coherence intervals can also be derived for inferences with more than one premise. The coherence interval for MP follows from the Equation and the total probability theorem of probability theory. Consider (4) again, inferring $E$ from *if L then E* and $L$. By the total probability theorem,

(6) $P(E) = P(L)P(E \mid L) + P(\textit{not-}L)P(E \mid \textit{not-}L)$.

When we definitely decide to take the long countryside route home, we judge $P(L) = 1$, and we can infer from (6) that our confidence that we will have an enjoyable journey is $P(E \mid L)$, which we might assess as .8 since there is some modest probability of a disabling condition that could make our journey unpleasant. Other people might agree that $P(E \mid L) = .8$, but not knowing what we have decided, perhaps set $P(L) = .5$. Using (6), they can then compute $P(L)P(E \mid L) = (.5)(.8) = .4$. When $P(L) = .5$, $P(\textit{not-}L)$ is of course also .5. The last part of equation (6), $P(E \mid \textit{not-}L)$, might be unknown. But by probability theory, it has to lie between 0 and 1. When it is 0, $P(E) = P(L)P(E \mid L) = .4$. And when it is 1, $P(E) = P(L)P(E \mid L) + P(\textit{not-}L) = .9$. To be coherent, the probability of the conclusion of MP for these people must then lie in the interval $[.4, .9]$, that is, between .4 and .9.

The derivations of coherence intervals for inferences (Evans, Thompson, & Over, 2015; chapter 4.4 by Pfeifer, this handbook; Pfeifer & Kleiter, 2009) yield a normative theory for the rational degrees of confidence people should have in the conclusions they infer from uncertain premises. To fall outside the coherence interval for an inference is to be incoherent and exposed to a *Dutch book* (i.e., a bet that will be lost no matter which possible outcome occurs). Dutch book arguments are the foundation of the normative Bayesian theory of coherence and rationality (Howson & Urbach, 2005).

People are not always coherent in their degrees of belief, but it is quite plausible that explicit inference would help them to be more rational at this basic level, for otherwise what benefit can come from the time and effort that go into explicit inference? Studies have shown that people are sometimes more coherent in the degrees of belief that they have in the conclusions of explicit inferences, but not necessarily (Cruz, Baratgin, Oaksford, & Over, 2015; Cruz, Over, Oaksford, & Baratgin, 2016; Cruz et al., 2017; Evans et al., 2015; Singmann et al., 2014). A striking example of when explicit inference does not facilitate coherence is a context in which the conjunction fallacy occurs (Tversky & Kahneman, 1983). In this context, people can incoherently judge $P(B) < P(A \& B)$ even when they explicitly infer $B$ from $A \& B$ (Cruz et al., 2015) and even though their judgments for the same inference are coherent when using neutral contents (Cruz et al., 2015; Politzer & Baratgin, 2016). This finding underlines the robustness of the fallacy and the importance of accounting for it (Costello & Watts, 2014; Tentori, Crupi, & Russo, 2013).

A further important inference form for the psychology of reasoning is *centering*. One-premise centering is sometimes called *conjunctive sufficiency*: it is the inference from $A \& B$ as a single premise to *if A then B* as a conclusion. Two-premise centering is the inference from $A$ and $B$ as separate premises to *if A then B* as a conclusion. Centering is a valid inference for a wide range of conditionals, including the material conditional, Lewis-style conditionals, and the suppositional conditional, and there is evidence that people conform to it in their degrees of belief (Cruz et al., 2015; Cruz et al., 2016; Pfeifer & Tulkki, 2017; Politzer & Baratgin, 2016). In contrast, the inference is invalid in *inferentialist* theories of the conditional (Douven, 2015, 2016; Douven, Elqayam, Singmann, & van Wijnbergen-Huitink, 2018). In semantic versions of these theories, *if A then B* means that there is a deductive, inductive, or abductive link between $A$ and $B$, and *if A then B* can only be true when there is such a link between $A$ and $B$. Supporters of this position argue that centering will be rejected for *missing-link* conditionals. These are conditionals *if A then B* with no relation between $A$ and $B$, and the claim is that these conditionals do not follow from the truth of $A \& B$. The following is an example of a missing-link conditional:

(7) If it is sunny in Canberra ($C$) today, then it is raining in London ($N$).

Supposing that $C$ and $N$ are both true, and so $C \& N$ is true, inferentialists would argue that (7) does not follow. It certainly could be pragmatically very odd in a discussion with someone to begin with $C$ and $N$ as premises, for the purpose of explicitly inferring (7) as a conclusion. On the other hand, if participants in an experiment are given a missing-link conditional *if A then B* first, followed by confirmation of $A$ and $B$, they hold that *if A then B* is true (Skovgaard-Olsen, Kellen, Krahl, & Klauer, 2017, figure 1). Indeed, if we were to assert (7) first in a discussion, however irrelevantly or oddly, and were to place a bet on it, we would surely claim to have stated a truth and to have won our bet when $C$ and $L$ turned out later to be true. This claim is in accordance with the de Finetti and Jeffrey tables for *if A then B* and a bet on it, and it has been strongly supported empirically in truth table tasks (see also Evans et al., 2003; Oberauer & Wilhelm, 2003), but it goes against the assumptions of inferentialism.

On the other hand, it is hard to see how (7), or any other missing-link conditional, can be used in rational decision making. When there is no epistemic relation between $A$ and $B$, making $A$ and $B$ independent, $P(B|A) = P(B)$, and $A$ cannot be a reason for doing $B$. There is evidence that in some contexts, people conform to the Equation more often when $P(B|A) > P(B|\textit{not-}A)$, suggesting an interpretation of *if A then B* in which $A$ raises the probability of $B$ (Skovgaard-Olsen, Singmann, & Klauer, 2016), but there is also evidence *against* this prediction (Oberauer, Weidenfeld, & Fischer, 2007; Over et al., 2007; Singmann et al., 2014). Cruz et al. (2016) find that the lack of an epistemic relation between $A$ and $B$ affects probability judgments not only about conditionals but also about conjunctions $A \& B$ and disjunctions *A or B*. This implies that a missing-link effect on probability judgments is not specific to conditionals, suggesting that it is pragmatic and not semantic. An effect that is not specific to conditionals cannot be used to distinguish the meaning of conditionals from that of other statement types. The findings of Cruz et al. (2016) also suggest that having a common topic of discourse between $A$ and $B$ is more important for the pragmatic acceptability of statements than a specific epistemic relation (see, on this, Over & Cruz, 2018; for a related account of why the computation of subjective probabilities for compound events is more difficult when $A$ and $B$ come from different domains, see Sanborn & Chater, 2016). The uselessness of missing-link conditionals for rational decision making does not have to be a matter of semantic meaning.

## 5. Rational Belief Updating

Suppositional reasoning is fundamental to rational decision making and to the belief revision and updating that goes with it. In scientific Bayesian inference, we might assign a low prior probability to a hypothesis $H$ at time 1, making $P_1(H)$ low, and infer (via Bayes' theorem) a high probability of $H$ supposing we get an finding $F$ in an experiment, making $P_1(H|F)$ high. Assume we now run an experiment at time 2 and learn that $F$ holds (and nothing else relevant). We can then use the new information about $F$ to update our degree of belief in $H$ at time 2, $P_2(H) = P_1(H|F)$, in an inference termed (*strict*) *conditionalization* (Howson & Urbach, 2005).

In a suppositional theory, we can identify this Bayesian process with conditional reasoning over time, in which we begin with $P_1(H)$ low and $P_1(\textit{if F then H})$ high, and then after we run the experiment and find that $P_2(F)$ holds, we can infer that $P_2(H)$ is high using *dynamic MP* (Oaksford & Chater, 2013): inferring high confidence in $H$ from high confidence in *if F then H* after learning $F$ as a result of acquiring new information. Bayesian inference like this is often an essential part of ordinary rational decision making.

A little more formally, suppose we begin at time 1 with a degree of confidence in $H$, say $P_1(H) = .5$, and in $H$ given $F$, say $P_1(H|F) = P(\textit{if F then H}) = .8$. At a later time 2, after acquiring new information, we judge that $P_2(F) = 1$, and we can then infer that $P_2(H) = .8$. When we are less sure of $F$ at time 2, say $P_2(F) = .7$, we can use

*Jeffrey conditionalization* (Jeffrey, 1983) to infer what our confidence in *H* should be at time 2:

(8) $P_2(H) = P_2(F)P_1(H \mid F) + P_2(not\text{-}F)P_1(H \mid not\text{-}F)$.

Jeffrey conditionalization is derived from the total probability theorem, an instance of which is (6). Both it and (strict) conditionalization depend on the assumption that the conditional probabilities are *invariant* (Oaksford & Chater, 2013), which means that $P_1(H|F) = P_2(H|F)$ and $P_1(H|not\text{-}F) = P_2(H|not\text{-}F)$. People do not always conform to Jeffrey conditionalization (Hadjichristidis, Sloman, & Over, 2014; Zhao & Osherson, 2010), perhaps because they tend to focus on the component—either $P_2(F)P_1(H|F)$ or $P_2(not\text{-}F)P_1(H|not\text{-}F)$—that has the higher probability and do not take into account the less probable one.

The problem of making a precise decision about $P_1(H|not\text{-}F)$ can be bypassed by considering the coherence interval: the minimum value of $P_1(H|not\text{-}F)$ is 0, and the maximum value is 1, allowing us to derive, as above, a coherence interval for Jeffrey conditionalization, justified when invariance holds. How far people's dynamic reasoning and belief updating are coherent, and their subjective conditional probabilities invariant, are important questions for future research.

## 6. Conclusion

In suppositional theory, a natural-language indicative conditional, *if A then B*, is equivalent to *B, supposing A*, and its probability is the subjective conditional probability of its consequent given its antecedent: $P(\text{if } A \text{ then } B) = P(B|A)$. This theory is a central part of a fully Bayesian approach to the psychology of reasoning and of judgment and decision making, which aims to unify these two subject areas and to study how degrees of belief are revised and updated in human reasoning, as well as how this process affects human decision making. The underlying standard of rationality in this approach, for both reasoning and decision making, is coherence, which means consistency with probability theory. People are sometimes coherent and sometimes incoherent in their reasoning and decision making, and unifying the two subject areas should lead to deeper understanding of why this is so.

### References


Adams, E. W. (1998). *A primer of probability logic*. Stanford, CA: CSLI Publications.

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

Baratgin, J., Over, D. E., & Politzer, G. (2013). Uncertainty and de Finetti tables. *Thinking & Reasoning*, *19*, 308–328.

Costello, F., & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychological Review*, *121*(3), 463–480.

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

Cruz, N., & Oberauer, K. (2014). Comparing the meanings of "if" and "all." *Memory & Cognition*, *42*, 1345–1356.

Cruz, N., Over, D., & Oaksford, M. (2017). The elusive oddness of or-introduction. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 259–264). Austin, TX: Cognitive Science Society.

Cruz, N., Over, D., Oaksford, M., & Baratgin, J. (2016). Centering and the meaning of conditionals. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1104–1109). Austin, TX: Cognitive Science Society.

de Finetti, B. (1995). The logic of probability (R. B. Angell, Trans.). *Philosophical Studies*, *77*, 181–190. (Original work published 1936)

de Finetti, B. (1964). Foresight: Its logical laws, its subjective sources. In H. E. Kyburg & H. E. Smokler (Eds.), *Studies in subjective probability* (pp. 55–118). New York, NY: Wiley. (Original work published 1937)

Douven, I. (2015). How to account for the oddness of missing-link conditionals. *Synthese*, *194*(5), 1–14.

Douven, I. (2016). *The epistemology of indicative conditionals*. Cambridge, England: Cambridge University Press.

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

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

Edgington, D. (2014). Indicative conditionals. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/win2014/entries/conditionals/

Elqayam, S., & Over, D. E. (2013). New paradigm psychology of reasoning: An introduction to the special issue edited by Elqayam, Bonnefon, and Over. *Thinking & Reasoning*, *19*, 249–265.

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

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. (2004). *If*. Oxford, England: Oxford University Press.

Evans, J. St. B. T., & Over, D. E. (2013). Reasoning to and from belief: Deduction and induction are still distinct. *Thinking & Reasoning*, *19*, 267–283.

Evans, J. St. B. T., Thompson, V. A., & Over, D. E. (2015). Uncertain deduction and conditional reasoning. *Frontiers in Psychology*, *6*, 398.

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*, 635–648.

Gilio, A., & Over, D. E. (2012). The psychology of inferring conditionals from disjunctions: A probabilistic study. *Journal of Mathematical Psychology*, *56*, 118–131.

Hadjichristidis, C., Sloman, S. A., & Over, D. E. (2014). Categorical induction from uncertain premises: Jeffrey's doesn't completely rule. *Thinking & Reasoning*, *20*(4), 405–431.

Howson, C., & Urbach, P. (2005). *Scientific reasoning: The Bayesian approach* (3rd ed.). Chicago, IL: Open Court.

Jeffrey, R. C. (1983). *The logic of decision* (2nd ed.). Chicago, IL: University of Chicago Press.

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

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove, England: Erlbaum.

Johnson-Laird, P. N., Khemlani, S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Science*, *19*, 201–214.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–156.

Lewis, D. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review*, *85*, 297–315.

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*, 346–379.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, *71*, 305–330.

Oaksford, M., Over, D. E., & Cruz, N. (2019). Paradigms, possibilities, and probabilities: Comment on Hinterecker et al. (2016). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*, 288–297.

Oberauer, K., Weidenfeld, A., & Fischer, K. (2007). What makes us believe a conditional? The roles of covariation and causality. *Thinking & Reasoning*, *13*(4), 340–369.

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*, 680–693.

Over, D. E. (2017). Causation and the probability of causal conditionals. In M. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 307–325). Oxford, England: Oxford University Press.

Over, D. E. (2020). The development of the new paradigm in the psychology of reasoning. In S. Elqayam, I. Douven, J. St. B. T. Evans, & N. Cruz (Eds.), *Logic and uncertainty in the human mind* (pp. 243–263). Abingdon, England: Routledge.

Over, D. E., & Baratgin, J. (2017). The "defective" truth table: Its past, present, and future. In N. Galbraith, D. E. Over, & E. Lucas (Eds.), *The thinking mind: The use of thinking in everyday life* (pp. 15–28). Hove, England: Psychology Press.

Over, D. E., & Cruz, N. (2018). Probabilistic accounts of conditional reasoning. In L. J. Ball & V. A. Thompson (Eds.), *International handbook of thinking and reasoning* (pp. 434–450). Hove, England: Psychology Press.

Over, D. E., & Cruz, N. (2019). Philosophy and the psychology of conditional reasoning. In A. Aberdein & M. Inglis (Eds.), *Advances in experimental philosophy of logic and mathematics* (pp. 225–249). London, England: Bloomsbury Academic.

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*, 62–97.

Pearl, J. (2013). Structural counterfactuals: A brief introduction. *Cognitive Science*, *37*, 977–985.

Pfeifer, N., & Kleiter, G. D. (2005). Coherence and nonmonotonicity in human reasoning, *Synthese*, *146*(1–2), 93–109.

Pfeifer, N., & Kleiter, G. D. (2009). Framing human inference by coherence based probability logic. *Journal of Applied Logic*, *7*, 206–217.

Pfeifer, N., & Kleiter, G. D. (2011). Uncertain deductive reasoning. In K. Manktelow, D. E. Over, & S. Elqayam (Eds.), *The science of reason: A Festschrift for Jonathan St B. T. Evans* (pp. 145–166). Hove, England: Psychology Press.

Pfeifer, N., & Tulkki, L. (2017). Conditionals, counterfactuals, and rational reasoning: An experimental study on basic principles. *Minds and Machines*, *27*(1), 119–165.

Politzer, G., & Baratgin, J. (2016). Deductive schemas with uncertain premises using qualitative probability expressions. *Thinking & Reasoning*, *22*, 78–98.

Politzer, G., Over, D. E., & Baratgin, J. (2010). Betting on conditionals. *Thinking & Reasoning*, *16*, 172–197.

Ramsey, F. P. (1990). General propositions and causality (original manuscript, 1929). In D. H. Mellor (Ed.), *Philosophical papers* (pp. 145–163). Cambridge, England: Cambridge University Press.

Ramsey, F. P. (1990). Truth and probability (original manuscript, 1926). In D. H. Mellor (Ed.), *Philosophical papers* (pp. 52–94). Cambridge, England: Cambridge University Press.

Sanborn, A. N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*, *20*(12), 883–893.

Singmann, H., Klauer, K. C., & Over, D. (2014). New normative standards of conditional reasoning and the dual-source model. *Frontiers in Psychology*, *5*, 16.

Skovgaard-Olsen, N., Kellen, D., Krahl, H., & Klauer, K. C. (2017). Relevance differently affects the truth, acceptability, and probability evaluations of "and," "but," "therefore," and "if–then." *Thinking & Reasoning*, *23*, 449–482.

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

Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98–112). Oxford, England: Blackwell.

Tentori, K., Crupi, V., & Russo, S. (2013). On the determinants of the conjunction fallacy: Probability versus inductive confirmation. *Journal of Experimental Psychology: General*, *142*, 235–255.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–315.

Zhao, J., & Osherson, D. (2010). Updating beliefs in light of uncertain evidence: Descriptive assessment of Jeffrey's rule. *Thinking & Reasoning*, *16*(4), 288–307.

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

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