American Psychological Association

Journal of Experimental Psychology: Learning, Memory, and Cognition

ISSN: 0278-7393

2020, Vol. 46, No. 4, 760–780
http://dx.doi.org/10.1037/xlm0000756

If and *or*: Real and Counterfactual Possibilities in Their Truth and Probability

Ruth M. J. Byrne
Trinity College Dublin, University of Dublin

P. N. Johnson-Laird
Princeton University and New York University

The theory of mental models postulates that conditionals and disjunctions refer to possibilities, real or counterfactual. Factual conditionals, for example, "If there's an apple, there's a pear," parallel counterfactual ones, for example, "If there had been an apple, there would have been a pear." A similar parallel underlies disjunctions. Individuals estimate the probabilities of conditionals by adjusting the probability of their *then*-clauses according to the effects of their *if*-clauses, and the probabilities of disjunctions by a rough average of the probabilities of their disjuncts. Hence, the theory predicts that estimates of the joint probabilities of these assertions with each of the four cases in their partitions will be grossly subadditive, summing to over 100%. Five experiments corroborated these predictions. Factual conditionals and disjunctions were judged true in the same cases as their counterfactual equivalents, and the sum of their joint probabilities with cases in the partition ranged from 240% to 270% (Experiments 1a, 1b). When participants were told these probabilities should not sum to more than 100%, estimates of the probability of *A and C*, as the model theory predicts, were higher for factual than counterfactual conditionals, whereas estimates of the probability of *not-A and not-C* had the opposite difference (Experiment 1c). Judgments of truth or falsity distinguished between conditionals that were certain and those that might have counterexamples (Experiment 2a), whereas judgments of the likelihood of truth reflected the probabilities of counterexamples (Experiment 2b). We discuss implications for alternative theories based on standard logic, suppositions, probabilistic logic, and causal Bayes networks.

Keywords: counterfactual conditionals, disjunctions, mental models, possibilities, probability

Supplemental materials: http://dx.doi.org/10.1037/xlm0000756.supp

Consider these two assertions:

If there is an apple, then there is a pear.

There isn't an apple and there isn't a pear.

Could they both—the conditional and the conjunction—be true at the same time? If so, what's their joint probability (from 0% to 100%)? Now, consider the following pair of assertions, where the conditional makes a counterfactual claim:

If there had been an apple, then there would have been a pear.

There isn't an apple and there isn't a pear.

Could they both be true, and what's their joint probability? If you know that there might be an apple without a pear, do you think that the two conditionals above are true, false, or impossible to determine?

The answers to these questions reveal the possibilities to which factual and counterfactual assertions refer. They are the focus of the present article. It begins with an outline of the theory of mental models—the model theory, for short—as it applies to compound assertions, such as the conditionals above, and disjunctions, such as "There is an apple in the bowl or there is a pear in the bowl, or both." The theory bases the meanings of such assertions on *possibilities*, which in turn underlie their probabilities: Numerical probabilities are possibilities plus numbers. After a summary of the theory and its three principal predictions, the article outlines alternative theories (see Nickerson, 2015)—those based on standard logic, suppositions, probabilistic logic, and causal Bayesian networks. It then describes five experiments that tested the three predictions. Finally, it considers the implications of the results for the model theory and for the alternative theories.

This article was published Online First October 24, 2019.

Ruth M. J. Byrne, School of Psychology and Institute of Neuroscience, Trinity College Dublin, University of Dublin; P. N. Johnson-Laird, Department of Psychology, Princeton University, and Department of Psychology, New York University.

The dataset for all of the experiments is available at https://osf.io/be9ck/ and https://reasoningandimagination.com/data-archive/. The research was supported by a grant from the Trinity College Dublin Arts and Social Sciences Benefactions Fund. We thank Shane Timmons for his help with Experiment 2a and Mark Keane, Nicole Cruz, Geoff Goodwin, Sunny Khemlani, Mike Oaksford, David Over, Cristina Quelhas, and Marco Ragni for their helpful comments. The authors have no declarations of interest.

Correspondence concerning this article should be addressed to Ruth M. J. Byrne, School of Psychology and Institute of Neuroscience, Trinity College Dublin, University of Dublin, College Green, Dublin, Ireland. E-mail: rmbyrne@tcd.ie

The Model Theory

The Meanings of Disjunctions

The model theory postulates that people construct models that simulate the situations that assertions describe (e.g., Johnson-Laird, 1983). Models are iconic in that they have a structure that corresponds to the structure of what they represent (e.g., Johnson-Laird, 2006). The theory provides a general account of reasoning, but here we describe only its explanation of sentential connectives, such as *if* and *or*. It treats such compound assertions as referring to real possibilities or counterfactual possibilities, which are those that were once real but that did not in fact happen (e.g., Byrne, 2005, 2016; Johnson-Laird & Byrne, 1991, p. 68). It distinguishes between their representations in mental models, which underlie intuitive reasoning (in System 1), and in fully explicit models, which underlie deliberations (in System 2). A factual disjunction, such as:

There is beer, or there is wine, or both

has three mental models (in System 1), which represent only the clauses that are true in each possibility:

beer

wine

beer wine

The fully explicit models of the disjunction (in System 2) represent in addition what is false in each possibility, using negation to do so:

beer not-wine

not-beer wine

beer wine

These three possibilities are exhaustive, and so they imply that given the disjunction the fourth case is impossible:

not-beer not-wine

The four cases based on the affirmation or negation of each clause in a compound assertion are known as its "partition'" *A & C*, *A & not-C*, *Not-A & C,* and *Not-A & not-C*, in which "&" denotes a conjunction in which *A & C* is equivalent to C *& A*. Of course, mental models are not words or sentences, which we use for convenience. They are actual models of the relevant entities or events. The two systems of the theory are both implemented in the computer program mSentential at https://mentalmodels.princeton.edu/models/.

An earlier version of the theory proposed that models were disjunctive alternatives—it is possible that there is beer alone, *or* it is possible that there is wine alone, *or* it is possible that there are both (e.g., Johnson-Laird & Byrne, 2002; Johnson-Laird, Byrne, & Schaeken, 1992). In a recent significant advance, however, the theory proposes that the models are exhaustive conjunctions of possibilities, which hold in default of knowledge to the contrary—it is possible that there is beer alone, *and* it is possible that there is wine alone, *and* it is possible that there is both beer and wine (e.g., Khemlani, Byrne, & Johnson-Laird, 2018):

possible (beer & not-wine)

& possible (not-beer & wine)

& possible (beer & wine)

Of course, in case there is neither beer nor wine the disjunction is false, i.e.,

impossible (not-beer & not-wine).

One sign that the possibilities are in a conjunction is that individuals accept each of the preceding possibilities as inferences from the disjunction (Hinterecker, Knauff, & Johnson-Laird, 2016), for example,

There is beer, or there is wine, or both.

Therefore, it is possible that there is beer.

The inference is invalid in all normal "modal" logics, which deal with possibilities and which exist in infinitely many sorts (Johnson-Laird & Ragni, 2019). Suppose that it is impossible that there is beer, but there is wine. The disjunction in the preceding inference is true in modal logics, but the conclusion is false. Hence, the inference is invalid in modal logics because in this example it leads from a true premise to a false conclusion. The model theory postulates that the inference can be drawn unless there is knowledge to the contrary, such as that it is impossible that there is beer. An inference is valid according to a semantic principle: there is no counterexample in which the premises are true but the conclusion is false (Jeffrey, 1981, p. 1). The model theory provides an algorithmic account of the cognitive processes that underlie the sentential inferences that people make, and it assesses the validity of a conclusion as necessary, possible, or impossible. Here, we focus on the descriptive claims of the theory rather than their normative status (see Johnson-Laird, Khemlani, & Goodwin, 2015; Johnson-Laird et al., 2019).

One consequence of the theory is that intuitive reasoning based on mental models can lead to compelling illusions, which deliberation with fully explicit models can correct. Consider the following example, which is based on two "exclusive" disjunctions, each referring to only two possibilities:

Either there's beer or else there's wine.

Either there isn't beer or else there's wine.

Can both of these assertions be true at the same time?

Most people say, "yes" (Johnson-Laird, Lotstein, & Byrne, 2012). The mental models of the two disjunctions are respectively:

beer

wine

and:

not-beer

wine

There is a model in common to the two disjunctions, which represents the presence of wine, and so individuals respond, "yes." However, the fully explicit models of the two disjunctions are:

beer not-wine

not-beer wine

and:

not-beer not-wine

beer wine

There is no model common to the two disjunctions—they are inconsistent with one another, and so they cannot both be true at the same time.

Models normally represent what is possible according to the truth of assertions. But, suppose a disjunction predicts the presence of wine or beer, or both, but in fact for some reason neither of them is there. The disjunction is false, but the situation can be described using a counterfactual disjunction:

There would have been wine, or there would have been beer, or both.

Its models represent the fact and counterfactual possibilities. A *counterfactual possibility* refers to something that once was possible but is not possible now. An *impossibility* is something that is not possible according to an assertion. The model theory deals with several distinct sorts of possibility (Johnson-Laird & Ragni, 2019), but here we confine our research to *epistemic* possibilities, which depend on knowledge, and are akin to subjective probabilities—Lassiter (2017) even treats the two as identical. Table 1 presents the possibilities and counterfactual possibilities—those that were once possible but that didn't happen—for factual and counterfactual conditionals and disjunctions.

The Meanings of Conditionals

As Table 1 illustrates, conditionals and disjunctions refer to possibilities. But, conditionals introduce a new factor. Unlike disjunctions that combine main clauses, the *if*-clause of a conditional is subordinate to its main *then*-clause. One sign of subordination is that pronouns in such clauses can refer forward (a "cataphor") to referents established in the main clause, for example, "If he served tea, then the butler wore white gloves." Subordinate clauses can make presuppositions. For example, the assertion "Before he left, he said goodbye" presupposes that he left. Presuppositions have to hold for an assertion to be true or false (see, e.g., Beaver & Geurts, 2014). But, an exception to this principle occurs with conditionals. According to the model theory, a conditional about a specific entity, such as a drink, for example,

If it's hot, then it's tea

asserts that it is possible that it is hot, and that in this case it is tea. And the *if*-clause presupposes the possibility that it is not hot. This presupposition fails in case it is impossible that it is not hot, that is, it is certain that it is hot, and so it follows from the conditional that it is tea. This analysis is general. Given *If A, then C*, the *if*-clause presupposes that it is possible that *not-A*, but in case the presupposition is false, it follows that *A* is certain.

This meaning for conditionals has several consequences. A conditional, *If A, then C*, has a conjunction of two mental models of possibilities:

A C

. . .

The first model makes explicit the possibility that the conditional asserts, and the second model, which the ellipsis denotes, has no explicit content but corresponds to the presupposition of the possibility of *not-A*. The fully explicit models of the conditional flesh out these models into a conjunction of three exhaustive possibilities that hold in default of knowledge to the contrary:

A C

not-A not-C

not-A C

Table 1
A Possibility Table for True Factual and Counterfactual Conditionals, and Disjunctions, About Specific Entities or Events, Showing the Modal Status of Each Case in the Partition as Possible, as Possible Once (But Did Not Happen), as Impossible, or as Factual, Given the Truth of the Assertion

Cases in the partition	*A & C*	*A & not-C*	*Not-A & C*	*Not-A & not-C*
A factual conditional *If A happened, then C happened*	Possible	Impossible	Possible	Possible
A counterfactual conditional *If A had happened, then C would have happened*	Possible once	Impossible	Possible once	Factual
A factual disjunction *A happened or C happened or both*	Possible	Possible	Possible	Impossible
A counterfactual disjunction *A would have happened or C would have happened or both*	Possible once	Possible once	Possible once	Factual

Note. "&" denotes a conjunction in which *A & C* is equivalent to *C & A*.

The models differ in their accessibility. The first of them is immediately accessible because it is explicit in the conditional's mental models, the second of them is next in accessibility because it holds whether the assertion is interpreted as a conditional or a biconditional (*If, and only if, A, then C*), and the third of them is least accessible because it holds only for a conditional interpretation. This trend in accessibility has been corroborated in many studies in which participants list possibilities (e.g., Barrouillet & Lecas, 1999). The conjunction of default possibilities distinguishes conditionals from the "truth functional" material conditionals of logic, which refer to a disjunction of true or false cases (see, e.g., Byrne & Johnson-Laird, 2009). Likewise, knowledge and meaning can modulate the interpretation of compounds by blocking the construction of models, or by adding temporal, spatial, and other relations, to their interpretations (e.g., Johnson-Laird & Byrne, 2002; Juhos, Quelhas, & Johnson-Laird, 2012; Quelhas & Johnson-Laird, 2017).

A well-known puzzle about conditionals is that people list the three possibilities and the one impossibility for *If A, then C*, but tend to judge that only two of them—those in which *A* holds—are relevant to the truth or falsity of conditionals. Hence, they judge the *not-A* cases in which the *if*-clause is false as irrelevant to the truth or falsity of conditionals (e.g., Evans, 1972). The puzzle is that a case that is possible for a conditional, such as *not-A and not-C*, ought to be one that people judge relevant to its truth value. The model theory solves the puzzle. The *not-A* possibilities are genuine, but presupposed, and so they are possible whether the conditional is true or false. They are therefore irrelevant to its truth value.

An unexpected consequence is that the probability of a conditional should equal the proportion of *A* possibilities in which *C* holds, because *not-A* possibilities are irrelevant to its probability too. Thus, it follows that the probability of *If A, then C* equals the conditional probability of *C* given *A*. This relation is a well-known Equation that probabilists defend (Edgington, 2014; Elqayam & Over, 2013; Oaksford & Chater, 2007; Pfeifer & Kleiter, 2010). As we currently show, however, naive individuals violate the probability calculus in a fundamental way.

To explain the model theory's account of counterfactual conditionals, we first need to refute an argument due to Adams (1970), which has led many theorists to suppose that factual and counterfactual conditionals must have quite different sorts of semantics. Lewis (1973), for instance, argued that counterfactuals require a "possible worlds" semantics, whereas factuals are the material conditionals of sentential logic (see also, Evans & Over, 2004; Jeffrey, 1981; Pearl, 2013). In Adams's argument, he contrasted the factual conditional:

> If Oswald didn't shoot Kennedy, then someone else did

with the counterfactual conditional:

> If Oswald hadn't shot Kennedy, then someone else would have.

The factual conditional is true, whereas the counterfactual is debatable. So, there is a radical difference between them. But, Adams's comparison is misleading. Any counterfactual in English has a parallel factual conditional, and the factual conditional that parallels the preceding counterfactual is:

> If Oswald hasn't shot Kennedy, then someone else will.

And its truth is just as debatable. Adams's factual conditional above takes for granted that someone shot Kennedy,[1] and so the parallel counterfactual should too. The past tense of "do" occurs in the factual conditional and does not cue a counterfactual interpretation, and so it is necessary to be more explicit:

> If Oswald hadn't been the person who shot Kennedy, then someone else must have been.

This counterfactual is true, just as the factual conditional is.

Adams (1975) considered counterfactuals as "epistemic past tense'" conditionals, that is, they have the same probability at the time of their utterance as factual conditionals on a prior occasion have. Hence, a factual conditional could have a low probability at the time at which the utterance of its counterfactual counterpart could have a high probability. He worried about this conjecture, because of his counterexamples in the previous paragraph. As we showed, however, his examples are not genuine parallels. Of course, factual and counterfactual conditionals do differ in meaning, but not in the radical way that Adams's argument implied (pace, e.g., Lewis, 1973).

Table 1 shows how a fact transforms a factual conditional into a counterfactual one. The fact matches a possibility, which becomes a factual case, the other possibilities become counterfactual possibilities, and the impossibility remains impossible. For false conditionals, the mappings swap the entries for the first two cases in their partition. An analogous transformation occurs from factual to counterfactual disjunctions. So, a systematic mapping exists from factual to counterfactual compound assertions. They run in parallel to one another (cf. Stalnaker, 1968, for such a treatment of conditionals).

Although the meanings of factual and counterfactual conditionals run in parallel, their *mental* models differ. A factual conditional, as we have seen, has only one mental model with explicit content, which represents the main possibility that the conditional asserts. A counterfactual of the sort, *If A had happened, then C would have happened*, has these mental models:

A	C	[possible once]
not-A	not-C	[fact]
. . .		

Hence, when someone asserts in an appropriate context "if there had been a circle . . . ," listeners know that the speaker has in mind that there is not a circle, and so the *if*-clause refers to a counterfactual possibility. Likewise, the *then*-clause, "there would have been a triangle" implies that there is not a triangle, but in the counterfactual possibility in which there is a circle, there is a triangle (Johnson-Laird & Byrne, 1991, pp. 68–69). Individuals try to keep track of the epistemic status of their models—whether they refer to facts, possibilities, or counterfactual possibilities. Counterfactual compounds can elicit knowledge of the actual status of the *if*-clause, the *then*-clause, and the relation between

[1] We thank David Over for pointing out in his review of a draft of this article that a similar observation that "Kennedy is dead" is true for the factual but uncommitted for the counterfactual is made by Pearl (2011).

them (Byrne & Tasso, 1999, p. 728; see also Byrne & Tasso, 1994, p. 126). They may know *not-A and not-C* as a matter of fact, or they may rely on the linguistic cue of the subjunctive mood to infer *not-A and not-C* is the case.

Many studies have corroborated this account (for a review, see Byrne, 2017). They show that tasks depending on the case of *not-A and not-C* are easier for counterfactual conditionals than for factual ones, as shown in the inferences people make (e.g., Byrne & Tasso, 1999), their judgments of what the conditionals imply (e.g., Thompson & Byrne, 2002; see also Fillenbaum, 1974), what they look at when they hear a counterfactual (e.g., Ferguson & Sanford, 2008), and the time they take to read the *not-A and not-C* case (e.g., Santamaría, Espino, & Byrne, 2005). Counterfactuals also represent the case of *A and C*, just as factual conditionals do (e.g., Santamaría et al., 2005). Hence, people readily make the modus ponens inference from counterfactual conditionals (Byrne & Tasso, 1994). Consider this example:

> If the burglar had been caught, then he would have been prosecuted.
>
> The burglar was caught.
>
> What follows?

It follows that the *if*-clause of the conditional is true, and the status of the conditional is updated to refer to the same possibilities as a factual one. Modus ponens follows at once: The burglar was prosecuted. Likewise, modus tollens is easier with a counterfactual than with a factual conditional, because the relevant mental model is now explicit. The phenomenon is robust, and inexplicable on other accounts.

The Probabilities of Compounds

Unlike theories founded on probabilities, the model theory implies that numerical probabilities enter into the contents of reasoning only if the task invokes them (Johnson-Laird et al., 2015, p. 207). So, when individuals judge whether a compound is true or false, the possibility of counterexamples is critical, but their probability only matters in tasks invoking probabilities. Likewise, a single counterexample can suppress inferences from factual conditionals (e.g., Byrne, 1989; Byrne, Espino, & Santamaría, 1999; see also De Neys, Schaeken, & d'Ydewalle, 2003), and from counterfactual conditionals (e.g., Espino & Byrne, 2019), but it is taken into account for ordinal judgments on a probabilistic scale (e.g., Markovits, Forgues, & Brunet, 2010; pace Geiger & Oberauer, 2007; Oberauer & Wilhelm, 2003). Likewise, a judgment of a conditional yields "false" when counterexamples exist, and it is not affected by whether their probability is high or low (Goodwin, 2014). The theory predicts, however, that this relative probability will be reflected in ordinal judgments of the probability of a conditional's truth.

Naive individuals are able to estimate the probabilities of many sorts of assertion. They can make simple inferences based on the frequencies or chances of an event, either from the proportions of equipossible models in which the event occurs, or from numerical tags on these models (Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999; Costello & Watts, 2016). But, they can also make inferences about the probabilities of unique events for which frequency data are not available, for example,

> The probability that the next Supreme Court justice is opposed to abortions is 90%.

They can use evidence and heuristics, as Tversky and Kahneman showed in their seminal research (e.g., Tversky & Kahneman, 1983). But, the mystery is this: Where do the numbers come from in such estimates?

A putative solution is based on the model theory (see Khemlani, Lotstein, & Johnson-Laird, 2012, 2015). Possibilities underlie probabilities: They are founded on the same semantics, but probabilities concern relative proportions or their numerical values (Johnson-Laird & Ragni, 2019). So, inferences from conditionals can refer to possibilities (as in Hinterecker et al., 2016), and probabilities can derive from them.

Estimates about unique events depend on the proportions of possibilities in models of relevant evidence. In the event that Trump nominates the next Supreme Court justice, the judge will be right wing, and the model theory postulates that individuals can adduce evidence:

> Most right-wing judges are opposed to abortion.

A mental model of the evidence yields a relevant proportion based on the quantifier, "most right-wing judges":

right-wing	opposed to abortion
right-wing	opposed to abortion
right-wing	opposed to abortion
right-wing	

Intuition (in System 1) represents this proportion in an iconic model of the subjective probability:

|------ |

where the left end of the scale represents impossibility, the right end represents certainty, and the length of the icon represents the probability. Various theories postulate that innumerate individuals represent magnitudes in such non-numerical models (e.g., Carey, 2009). A comparable claim follows from the model above:

> It's highly likely that the next Supreme Court justice is opposed to abortion.

Further evidence can shift the magnitude of the icon one way or another, and so it can represent the probability of compound assertions. A similar process occurs for conjectures, such as whether there is an apple in the fruit bowl, or a pear, or both. Individuals who know nothing about the situation are likely to assume that each possibility is equiprobable (for corroboration, see Johnson-Laird et al., 1999).

Most people do not know the correct ways to estimate the probabilities of compound assertions, and the model theory proposes that their intuitive estimates (in System 1) of the probability of a conjunction or a disjunction of *A* and *C* is a rough average of their estimates of the probability of *A* and the probability of *C*. Their degree of belief in the compound depends on their degrees of belief in its elements. Likewise, their intuitive estimate of the probability of *If A, then C* relies on making an estimate of the probability of *C*, and then nudging its value either higher or lower according to what effect, if any, they infer that *A* should have on the probability of *C*. To estimate the probability of a counterfactual conditional, *If A had happened, then C would have happened*, individuals have to envisage the point in time at which *A* was a real possibility and then rely on the same procedure as for the factual, *If A, then C*, but with a further factor. They now know that the probability of *A* is reduced, because it did not, in fact, happen. This algorithm seems close to an implementation of Adams's (1975) conception of an epistemic past tense.

The model theory explains why Adams's (1970) original examples about Kennedy's assassination differ so radically. Individuals know that someone shot Kennedy, and so they evaluate the factual conditional ("If Oswald didn't shoot Kennedy, then someone else did") in this context. They infer from someone shot Kennedy and it was not Oswald, that it was someone else. So, the conditional is true. For the counterfactual conditional ("If Oswald hadn't shot Kennedy, then someone else would have"), individuals envisage the point in time in which one possibility was that Oswald did not shoot Kennedy, and then consider any relevant evidence that someone else thereafter will shoot him. In the absence of any strong evidence for such an event, they infer that there are very few possible scenarios in which it will happen. If asked for a numerical estimate, they give a low probability.

The deliberations in System 2 are more sophisticated than those for intuitions: they multiply their estimates for a conjunction, and sum them for a disjunction, but they fail to consider the potential dependence of one event on the other. And, for conditionals, they estimate the proportion of cases of *A* in which *C* holds (a procedure equivalent to the Equation for conditional probabilities). The model theory of intuitive and deliberative probabilities is implemented in a computer program, *mReasoner*, which integrates deduction and probabilities (Khemlani et al., 2015). It uses loops of a small fixed number of iterations to form the intuitive averages of two values on an icon: It moves them toward each other until they meet. Deliberations, which are computationally more powerful, can map the resulting iconic model into a numerical estimate of a probability.

Experiments have corroborated the hypothesis that estimates of the probability of *A* and of the probability of *C* underlie estimates of the probability of a compound of *A* and *C*. Explicit estimates of these three probabilities fix the "joint probability distribution (JPD)," that is, the probabilities of each of the conjunctions in the partition: *A & C*, *A & not-C*, *not-A & C*, and *not-A & not-C*. But, the procedures described above for compounds should tend to yield what mathematicians refer to as subadditive values for the joint probability distribution, that is, they sum to more than 100%, contrary to the probability calculus. Square roots, for instance, are subadditive, for example, $\sqrt{(2 + 2)}$ is less than $\sqrt{2} + \sqrt{2}$, but probability in the standard calculus is not subadditive, because the probability of the complete joint probability distribution equals the sum of the probabilities of its parts—each of the cases in the partition. Moreover, granted that deliberations for conditionals correspond to Bayes's theorem, the model theory predicts that conditionals should not be as subadditive as disjunctions for which deliberations overlook the possible dependency between the two events. Experiments corroborated the prediction: Estimates for conditional probabilities and their two clauses were subadditive on only 24% of trials, whereas estimates for disjunctions and their two clauses were subadditive on 65% of trials (Khemlani et al., 2015). For an inclusive disjunction, *P(A or C)*, participants based their intuitive estimates on a primitive average of *P*(*A*) and *P*(*C*), but for a conditional probability, *P*(*C*|*A*), they treated *A* as evidence for estimates of *P*(*C*). Participants in our first experiment had to make a different sort of estimate: they estimated the probabilities of the conjunction of a compound assertion with each of the four cases in its partition. One such estimate, for example, was of the joint probability of

> If there is an apple, then there is a pear, and
>
> there isn't an apple and there isn't a pear.

The sum of the four estimates according to the standard probability calculus equals 100%.

The Model Theory's Predictions

The model theory yields three principal predictions about conditionals and disjunctions.

Prediction 1: Truth. Participants will judge that a compound and each of the cases in its partition that are possible can be jointly true. Judgments for disjunctions will not differ reliably over their possibilities, because mental models represent each of them, and so each of the three possibilities, *A & not-C*, *not-A & C*, *A & C*, will tend to be judged to be true equally often. In contrast, judgments for conditionals will show a reliable trend based on the accessibility of their possibilities, and so there will be a greater tendency to judge *A & C* to be true compared to *not-A & not-C*, which in turn will have a greater tendency to be judged true compared to *not-A & C*. Judgments of counterfactuals will parallel those of factuals, given that their fully explicit models are the same, differing only in their epistemic status.

Prediction 2: Probability. Participants' estimates of the joint probabilities of a compound and the cases in its partition will be grossly subadditive, because they are making four judgments that are each based on a process that tends to be subadditive. For example, they may make intuitive estimates that the probability of *if A, then C* and *A & C* is very high, the probability of *If A, then C* and *not-A & not-C* is also quite high, the probability of *If A, then C* and *not-A & C* is high, and the probability of *If A, then C* and *A & not-C* is zero. When they translate these intuitions into numerical chances out of 100, they may assign high numbers to reflect their judgments of probability, for example, 90 to *If A, then C* and *A & C*, 70 to *if A, then C* and *not-A & not-C*, and so on. As a result, their numerical estimates will exceed 100, in violation of the standard probability calculus. The prediction follows from the model theory's account of how individuals estimate the probabilities of compound assertions (Khemlani et al., 2015). Conditionals will yield less subadditivity than disjunctions, for which estimates overlook the possible dependency between the two events. Counterfactual conditionals will yield less subadditivity than factual conditionals, because one of the cases in their partition corresponds to the presupposed facts. When individuals know that the sum of their estimated probabilities of the four cases in a compounds' partition should not sum to more than 100%, a subtle interaction will occur. For conditionals, *A and C* will be judged more probable for factual than counterfactual conditionals, because it has an explicit mental model for factuals but is only a counterfactual possibility in the mental models for counterfactuals.

In contrast, *not-A and not-C* will be judged less probable for factual than counterfactual conditionals, because it has no mental model for factual conditionals but does for counterfactual conditionals for which it is a fact.

Prediction 3: Counterexamples. Judgments of the truth-values of compounds will reflect only the possibility of a counterexample, and so participants will tend to judge as true conditionals with no counterexamples more often than conditionals with counterexamples, regardless of the number of counterexamples. In contrast, their estimates of the probabilities of their truth-values will be sensitive to the probability of counterexamples, and so they will estimate the probability of truth as highest for conditionals with no counterexamples, as intermediate for conditionals with a few counterexamples, and as lowest for conditionals with many counterexamples.

Alternative Theories of Compounds

Four main alternative psychological theories of compound assertions exist. They are based on inference rules from standard logic, on suppositions, on probabilistic logic, and on causal Bayes networks. We consider each sort of theory in turn.

Theories Based on Logical Inference Rules

Theories based on mental proofs using formal inference rules from standard logic treat conditionals, *If A, then C*, as material conditionals, which are truth-functional, that is, they are true in any case in their partitions except for *A and not-C* in which case they are false (e.g., Rips, 1994). But, such interpretations yield "paradoxes," for example, conditionals are true provided that their *if*-clauses are false, so the following sort of inference is valid:

> The government is not going to allow fracking.
>
> Therefore, if the government is going to allow fracking, then it will lose the next election.

The inference is absurd. But, one way in which to try to save the material conditional, following Grice (1989), is to appeal to the conventions governing conversation. A speaker asserting the conditional above would not do so if she believed that the government was not going to allow fracking. So, the utterance conveys a conversational implicature that its *if*-clause is true. This implicature prevents the paradox above. A conversational implicature is not a valid deduction, and so it can be cancelled, as in this utterance:

> It's sensible of the government not to allow fracking, because it will lose the next election if it does.

Its first clause cancels the implicature that the subsequent *if*-clause is true. So, we are back to a material conditional in this example, and because its *if*-clause is false, the conditional itself is bound to be true, granted that the government is not going to allow fracking. The absurd consequence is a prediction about the next election that is bound to be true too. In sum, Grice's conversational implicatures cannot save the approach to conditionals based on standard logic. Such theories also cannot explain the difference in accessibility of the cases in a conditional's partition or the parallel between factual and counterfactual compounds, because factual compounds are "truth functional," and counterfactuals cannot be truth functional (see, e.g., Jeffrey, 1981, Ch. 4). We say no more about theories based on logical inference rules.

The Suppositional Theory of Conditionals

The suppositional theory of conditionals postulates that *if* triggers a supposition of its clause, and individuals evaluate the *then*-clause in this hypothetical situation (Evans & Over, 2004). The theory rejects the material conditional of logic, and instead postulates that conditionals have no truth value in the cases in which their *if*-clauses are false. People therefore understand a factual conditional by adding its *if*-clause to their beliefs and calculating the probability of its *then*-clause. They think about true *if*-clauses only and not about their falsity or whether or not it would imply the *then*-clause (e.g., Handley, Evans, & Thompson, 2006; see also Oaksford & Chater, 2007). In general, proponents of suppositions have more recently adopted a probabilistic approach to conditionals (see, e.g., Evans, 2012; Over, 2009) to which we now turn.

Probabilistic Logic

Various theories of conditionals are based on Adams's (1998) probabilistic conditional (see, e.g., Liu, Lo, & Wu, 1996; Oaksford & Chater, 2007; Oaksford, Chater, & Larkin, 2000; Over, Hadjichristidis, Evans, Handley, & Sloman, 2007; Pfeifer & Kleiter, 2005, 2010). On this account, "an ordinary conditional assertion if p then q is interpreted as q is probable given p" (Evans, Handley, & Over, 2003, p. 322). Hence, people believe a conditional if its corresponding conditional probability is high enough to warrant the judgment (Oberauer & Wilhelm, 2003, p. 685). As we saw earlier, the model theory also implies this Equation for the subjective probability of a conditional, but not for its meaning.

Over and Cruz (2017) remarked that probabilists have neglected counterfactuals, but they have noted commonalities between factual and counterfactual conditionals (e.g., Elqayam & Over, 2013; Evans & Over, 2004; Over et al., 2007). Adams (1975) proposed that the probability of a counterfactual at the present time is the same as the probability of the corresponding factual conditional at a previous time. And to understand a counterfactual is to estimate its believability. Such an estimate depends on its conditional probability, that is, the likelihood of its *then*-clause given its *if*-clause. Individuals think about only one situation at a time (the singularity principle) and so the counterfactual is represented in a single mental representation (Evans, 2007, p. 74). The theory allows that individuals can represent an implicature of the facts *A* and *C* (Evans, Over, & Handley, 2005, p. 1049). However, a significant modification to Adams's (1998) account of factual conditionals is that the probability of a counterfactual, *If A had happened, C would have happened*, combines its corresponding conditional probability with the degrees of belief in each of the implicated facts, *A* and *C* (Elqayam & Over, 2013).

Few experimental studies have tested such probabilist ideas about counterfactuals (Over & Cruz, 2017). But, one study observed no reliable differences between indicative and counterfactual conditionals in participants' tendency to produce conclusions consistent with a conditional probability interpretation for various inferences such as modus ponens (Pfeifer & Tulkki, 2017). In another study (Over et al., 2007, Experiment 3), participants judged the probability of counterfactuals such as:

> If New York had not been attacked by terrorists in 2001, then the United States would not have attacked Iraq

and the probability of each of the four cases in its partition, but from a point in time prior to the events in 2001, for example,

> New York will not be attacked by terrorists and the United States will not attack Iraq.

The participants were told that the four probabilities should sum to 100%. Their estimates of the probability of a counterfactual, *If A had not happened, then C would not have happened*, correlated with the estimates of the earlier probability of *not-C* given *not-A*, which the experimenters computed from the estimates of the relevant cases in the partition. This result corroborates both the probabilist account and the model theory.

Otherwise, probabilist theories appear to have the following consequences. First, they make no predictions about disjunctions or about differences between them and conditionals. They make no predictions about differences between factual and counterfactual conditionals, because they make no claims about the possibilities to which compound assertions refer. Second, they predict that estimates of the joint probabilities of a compound and the cases in a compound's partition should tend to be coherent rather than to violate the probability calculus. Of course, if individuals have beliefs about the probability of various cases in the partition, they may estimate different probabilities for them, but nonetheless they should sum to 100 (Pfeifer & Kleiter, 2010). Third, probabilist theories predict that judgments of the truth values of compounds and estimates of the probability of their truth-values will both be sensitive to the probability of counterexamples. Conditionals with no counterexamples and those with a low probability of counterexamples will both tend to be judged to be true, because the meaning of a conditional tolerates exceptions, and both will tend to be judged to be true more often than conditionals with many counterexamples.

Causal Bayes Networks and Counterfactuals

An alternative probabilistic theory of counterfactual conditionals is due to Pearl (2009, 2011, 2013; see also Spirtes, Glymour, Scheines, et al., 2000). It treats inferences from factual conditionals and disjunctions in the same way as standard logic. It treats counterfactual conditionals as concerning causation (cf. Lewis, 1973, for an analogous distinction). Pearl's aim was to get scientists and statisticians to take causation seriously in structural equations, and to enable them to compute both deterministic causes and causal probabilities. His starting point was Bayes networks. They are a parsimonious way to represent interrelated conditional probabilities in place of the full joint probability distribution. In a network, an arrow from one variable to another stands for a direct relation, but a route from one variable to another may also be indirect. Each node in the network can be linked to a representation of the conditional probability of the variable's value given its inputs. Pearl's concept of causation is that it is a summary of what happens when an intervention establishes a counterfactual conditional. For example, suppose you want to determine whether daily exercise causes a reduction in blood pressure. You may observe a correlation, but a decisive test calls for an experiment. In daily life, whether or not a person does daily exercise depends on many vagaries, for example, the person is a former athlete, which in turn may also have a direct effect on blood pressure. So, a causal network will have a direct link from "former athlete" to "blood pressure," and an indirect link between the two via "exercises daily." An intervention severs these links in the network (using a so-called do-operator). In a corresponding actual experiment, the participants would exercise daily or not depending only on the condition in the experiment to which they were assigned. Once you have a causal network of this sort, you can assign a probability distribution to it, and Pearl (2009) gave a procedure by which to compute the probability of a counterfactual: "If a person had exercised daily then the person would have lowered blood pressure" (Ch. 7). Pearl's approach solves problems for Lewis's (1973) counterfactual approach, such as overdetermination in which an effect has multiple causes (see also Lagnado, Gerstenberg, & Zultan, 2013; Oaksford & Chater, 2017).

Sloman and Lagnado (2005) proposed that humans rely on such networks for causal reasoning. They argued that people can envisage an intervention using the do-operator, and set a variable to a particular value to see what the consequences are. Following Pearl (2009, Section 1.4.4), they propose a three-step procedure for answering a counterfactual question:

> If B had not occurred, would D still have occurred?

First, you infer what follows from the network using standard logic, for example, given *D* occurred, then its precursors *A*, *B*, and *C* also occurred. Second, the question refers to a counterfactual possibility in which *B* did not occur. So, you make an intervention and the do-operator prunes the causal links to *B*, and sets *B* to not having occurred. Third, you infer the consequences of your intervention, for example, in a network in which *A* causes *B* and *C*, which each cause *D*, then given that *A* occurred, *C* occurred, which suffices for *D* to occur. So, the answer to the question is "yes" (see also Sloman, 2013). Reasoning about interventions cannot be carried out in standard logic, but experiments corroborated the predictions. However, subsequent studies have shown that contrary to pruning, reasoners do think about causes upstream from the point of intervention (Rips, 2010; see also Rips & Edwards, 2013).

Individuals recognize the difference between observations and interventions, including counterfactual interventions for which the real state is known, hypothetical interventions for which the real state is not known (Lucas & Kemp, 2015; Meder, Hagmayer, & Waldmann, 2009), and interventions informed by explanations and norms (Dehghani, Iliev, & Kaufmann, 2010). They also use heuristics to determine the best intervention by identifying nodes as causes or effects of a target node (Meder, Gerstenberg, Hagmayer, & Waldmann, 2010). We return to causal Bayes networks in the General Discussion, but, they do not appear to make any of the model theory's three principal predictions: (1) that participants will judge the joint truth of conditionals and disjunctions with the possibilities to which they refer, (2) that their estimates of the probabilities of their joint truth will be more subadditive for disjunctions than for conditionals, and (3) that their judgments of the truth of compounds given counterexamples will be insensitive to their probability unlike their estimates of the probability of truth.

Experiment 1a

Our initial experiment tested the model theory's Predictions 1 and 2. It examined participants' judgments of the joint truth and joint probability of compound assertions and each of the cases in their partitions. The compounds were conditionals or disjunctions, and factual or counterfactual. For example, to test the first prediction for a counterfactual conditional,

> If there had been an apple in the fruit bowl, there would have been a pear

the participants judged whether its conjunctions with each case in its partition, for example,

> There was an apple in the fruit bowl and there was a pear

can both be true. This task should not be confused with one in which participants select potential evidence bearing on the truth or falsity of a conditional, which leads them to treat cases in which the *if*-clause is false as irrelevant—a sensible reaction according to the model theory. In contrast, the present task is in effect asking whether the pairs of sentences are consistent with one another. (We discovered the hard way that framing the task in terms of "consistency" confuses participants, and so the experiment frames it in an equivalent way: "can both the assertions be true?"; see Khemlani, Orenes, & Johnson-Laird, 2014). To test the second prediction, the experiment calls for them to then estimate the chances out of 100 that both assertions are true. The model theory predicts that a conditional can be jointly true with any case in its partition except *A & not-C*, but the other cases should yield the following trend in acceptance: *A & C*, *not-A & not-C*, and *not-A & C*. An inclusive disjunction, *A or C, or both*, can be jointly true with any case in its partition except *not-A & not-C*, and its possibilities should be equally accessible. According to the model theory's second prediction, estimates of the joint probability of a compound and the cases in its partition should be highly subadditive, greater for disjunctions than conditionals, and greater for factual compounds than counterfactual compounds. To test this prediction, we chose materials, such as "if there was an apple in the fruit bowl then there was a pear" for which participants would not have strong beliefs on which to base their judgments.

Method

Participants. The participants were 94 volunteers recruited from the online platform Prolific (www.prolific.ac), restricted to English speaking countries. They were paid £1.25 sterling for their participation in the experiment, which took about 15–20 min. For all the studies, the participants gave their informed consent, and we report all our manipulations and measures, and each study's sample size was determined prior to data collection. We had planned a sample size of 80 participants to ensure sufficient power to observe a small to medium effect and set the preprogrammed recruitment stopping rule to recruit until 100 participants had been assigned by the program (on the assumption that some participants would fail the attention checks). The materials were presented using SurveyGizmo software (www.surveygizmo.com). Prior to analysis, we eliminated the data from eight participants: two because they failed questions commonly used in online experiments to check that they had paid attention and six because they made bizarre estimates of probability, for example, negative numbers, or estimates of over 400 on at least one trial. As a result, the experiment tested 72 women and 22 men, ranging in age from 19 to 71 years with a mean of 39 years. All of the experiments received ethical approval from the School of Psychology Ethics Committee at Trinity College Dublin.

Design, materials, and procedure. The participants carried out eight trials, two for each of four sorts of compound assertion:

1. Factual conditionals, for example, "If there was an apple in the fruit bowl, then there was a pear."
2. Counterfactual conditionals, for example, "If there had been an apple in the fruit bowl, then there would have been a pear."
3. Factual disjunctions, for example, "There was an apple in the fruit bowl or there was a pear, or both."
4. Counterfactual disjunctions, for example, "There would have been an apple in the fruit bowl or there would have been a pear, or both."

We devised eight sets of contents, which were about flowers, fruit, vegetables, and other foods, which were in vases, bowls, and other receptacles (see the online supplemental materials). The eight contents were assigned at random to the eight problem forms. The problems were presented to each participant in a random order.

The participants were instructed to take part in the study only if they were prepared to consider the descriptions seriously and to give some thought to answering the questions carefully. Their task is illustrated in the following example of a trial for a counterfactual conditional (adapted from Goodwin & Johnson-Laird, 2018):

> Consider the following *if–then* statement:
>
> > If there had been a rose in the vase, then there would have been a daffodil.
>
> Now consider each of the four *and* statements on the left below. For each statement, please indicate whether it and the *if–then* statement can *both be true*.
>
> Alongside that judgment, please also indicate your estimate of the *chances* that both are true. In the case that you judge that a pair of statements cannot both be true, then plainly the chances that they are both true must be 0.
>
> The *if–then* statement is repeated here:
>
> > If there had been a rose in the vase, then there would have been a daffodil.

	Could both this statement and the *if–then* statement above *both be true*?	What are the chances that both this statement and the *if–then* statement above are *both true*?
	Write "yes" or "no"	Write a number from 0 (*no chance at all*) to 100 (*completely certain*).
There was a rose and there was a daffodil.	_____	_____
There was a rose and there was *no* daffodil.	_____	_____
There was *no* rose and there was a daffodil.	_____	_____
There was *no* rose and there was *no* daffodil.	_____	_____

As the example illustrates, the participants made four judgments of joint truth and four judgments of joint probability for each compound and the cases in its partition, that is, a total of 64 judgments. They were given two practice problems based on a factual conditional and a factual disjunction in which they made judgments only of joint truths.

Results and Discussion

The dataset for all of the experiments is available at https://osf.io/be9ck/ and https://reasoningandimagination.com/data-archive/. Table 2 presents a summary of the overall proportions of judgments that a compound and each case in its partition could both be true. As the model theory predicts, the participants tended to judge that conditionals can be jointly true with the cases *A and C*, and *not-A and not-C*, and that they cannot be jointly true with the case *A and not-C*. These judgments do not reflect a response bias because the judgment of joint truth was reliably higher for *A and C* than for *A and not-C* (Wilcoxon's test, $z = 8.61$, $p < .001$, $r = .63$). Their only judgments that were not reliably different from chance were for the joint truth of the conditionals with the case *not-A and C*. Some conditionals may have been interpreted as biconditionals of the sort, *If, and only if, A, then C*, for which the correct judgment is that this case cannot be jointly true. Likewise, the model theory predicts that the judgments of the four cases should be the same for both factual and counterfactual conditionals. They were identical on at least one of the two trials with factual and counterfactual conditionals for 77 out of the 94 participants, and identical on both trials for 55 of them (binomial test, with a prior of .25, given the four cases, $p < .001$).

Table 2
The Mean Proportions of Judgments in Experiment 1a That the Compound and Each Case in the Partition Can Both Be True

	Compound assertions			
	Conditionals		Disjunctions	
Truth judgments	Factual	Counterfactual	Factual	Counterfactual
Cases in the partition				
A and C	.96	.93	.97	.95
Not-A and not-C	.84	.86	.18	.13
Not-A and C	.49	.46	.79	.83
A and not-C	.10	.10	.79	.85

Note. All but two proportions (.46 and .49) in this table were either significantly greater than or significantly less than the chance value of .5 (Wilcoxon tests, $p < .001$, where Bonferroni corrected alpha is $p < .003$). See Table S1 in the online supplemental materials for standard deviations.

As Table 2 also shows, the participants tended to judge that disjunctions can be jointly true with the cases of *A and C*, *A and not-C*, and *not-A and C*, and that they could not be jointly true with cases of *not-A and not-C*. These judgments did not reflect a response bias, because the judgment of joint truth was reliably higher for the case of *A and C* than for the case of *not-A and not-C* (Wilcoxon's test, $z = 8.44$, $p < .001$, $r = .62$).

Table 3 summarizes the participants' mean estimates of the joint probability of a compound and each case in its partition. If these estimates are to be consistent with the probability calculus, they should never sum to more than 100. In fact, as the model theory predicts, the four sums were highly subadditive, ranging from 240 to 270 chances out of 100. Overall, 92 out of the 94 participants made subadditive estimates (binomial test, $p < 1$ in 10 million).

As Table 3 suggests, estimates of the probabilities for disjunctions were more subadditive than those for conditionals (Wilcoxon's test, $z = 3.86$, $p < .001$, $r = .28$), and those for factuals were more subadditive than those for counterfactuals (Wilcoxon's test, $z = 2.1$, $p < .036$, $r = .15$). These two variables did not interact reliably (Wilcoxon's test, $z = 1.4$, $p < .16$). But, factual conditionals were more subadditive than counterfactual conditionals (Wilcoxon's test, $z = 2.32$, $p < .026$, $r = .16$), whereas no reliable difference occurred between factual and counterfactual disjunctions (Wilcoxon's test, $z = .87$, $p < .385$). The complete set of statistical comparisons is with Table S1 in the online supplemental materials.

Overall, the results supported the model theory's first two predictions. The decline in the participants' tendency to endorse the joint truth of conditionals, *If A, then C*, over the cases *A and C*, *not-A and not-C*, *not-A and C*, corroborated the theory's first prediction, as did the endorsement of the disjunctions and their possible cases, which did not differ in reliability. The vast subadditivity of the participants' estimates of the joint probabilities of the compounds and cases in the partition corroborated the model theory's second prediction. So, too, did the greater subadditivity of disjunctions over conditionals, and the greater subadditivity of factual over counterfactual conditionals. Probability estimates of a single compound and of its constituent propositions tended to be subadditive (Khemlani et al., 2015), so the theory predicts that four such estimates should amplify the effect. Its magnitude, however, was massive.

A potential alternative to the model theory's prediction of subadditivity, which we owe to Mike Oaksford (personal communication, September 2018), is that the participants carried out a different task. They did not estimate the joint probability of, say, *If A, then C*, and *A and not-C*, but instead they estimated the conditional probability of *not-C* given *A*. Likewise, for the case in the partition of *A and C*, they estimated the conditional probability of *C* given *A*, and likewise for the other cases. This alternative hypothesis, however, runs into several difficulties. There is no obvious reason why the participants should convert the task into one calling for a conditional probability (see the presentation of a typical trial above). Moreover, when participants estimate the probability of *A*, the probability of *C*, and the conditional proba-

Table 3
Mean Estimates of the Joint Probability (From 0 to 100) in Experiment 1a of a Compound Assertion and Each Case in Its Partition

	Compound assertions			
	Conditionals		Disjunctions	
Probability estimates	Factual	Counterfactual	Factual	Counterfactual
Cases in the partition				
A and C	86	83	83	78
Not-A and not-C	72	72	49	45
Not-A and C	53	47	68	71
A and not-C	42	38	70	72
Sum	253	240	266	270

Note. The sum of probabilities for each of the four sorts of compound was reliably greater than the probability calculus's normatively correct sum of 100 (Wilcoxon tests in each case, $z \geq 8.2$, $p < .001$, $r = .60$.) See Table S1 in the online supplemental materials for standard deviations.

bility of one given the other, they are reliably subadditive (Khemlani et al., 2015), and the hypothesis offers no explanation for this result. Furthermore, the sum of the two preceding conditional probabilities ought to be no more than 100, but as Table 3 shows it was 128 chances out of 100. The hypothesis also fails to explain the vast subadditivity of the joint estimates with disjunctions: no plausible reason exists to suppose that disjunctions should trigger estimates of conditional probabilities of one disjunct given the other. Of course, the participants erred in judging impossible cases as capable of joint truth and in estimating them as having probabilities greater than zero. Yet, even if we discount these estimates, the remaining subadditivity was vast.

Experiment 1b

In case the gross subadditivity in the previous experiment was a result of a misunderstanding, this experiment both simplified the task and clarified the import of an estimate of 100 chances. The experiment called only for estimates of the probability of each case in a compound's partition, that is, the judgments were no longer joint with the compounds. The experiment also compared the old instructions for estimating probabilities with a new instruction that emphasized that an estimate of 100 implied that the case held in each and every possibility.

Method

Participants. The participants volunteered on the online platform Prolific and were paid £1.25 sterling for their participation. We eliminated the data of 17 potential participants, six because they had carried out a similar task, five because they failed the attention test, one for being underage, and the remainder for making the same estimate for all problems, or failing to answer all questions. There were 81 resulting participants: 60 women, 20 men, and one individual with a nonbinary gender, and their ages ranged from 20 to 73 years with a mean of 40 years.

Design, materials, and procedure. The design was based on the previous experiment but with two changes. First, the participants were assigned at random to one of two groups for which their instructions for making their probability estimates differed slightly. One group of participants ($n = 43$) was told to write a number from 0 (*no chance at all*) to 100 (*completely certain*). Another group of participants ($n = 38$) was told to write a number from 0 (*no chances at all*) to 100 (*all chances, each and every one*). The point of this second formulation was to ensure that the participants grasped the significance of an estimate of 100. Second, the participants no longer made joint judgments of the truth or probability of the compound together with a case in the partition. Instead, they made judgments only of each case in the partition, albeit in the context of a compound. An example of the task is as follows:

> Consider the following *if–then* statement:
>
> If there had been a rose in the vase, then there would have been a daffodil.
>
> Now consider each of the four *and* statements on the left below. For each statement, please indicate whether it can be *true*.
>
> Alongside that judgment, please also indicate your estimate of the *chances* that it is true. In the case that you judge that it cannot be true, then plainly the chances that it is true must be 0.
>
> The *if–then* statement is repeated here:
>
> If there had been a rose in the vase, then there would have been a daffodil.

	Could the statement on the left be TRUE? Write 'Yes' or 'No'	What are the chances that the statement on the left is TRUE? Write a number from 0- 'no chances at all' to 100- 'all chances, each and every one'.
There was a rose and there was a daffodil.	☐	☐
There was a rose and there was NO daffodil.	☐	☐
There was NO rose and there was a daffodil.	☐	☐
There was NO rose and there was NO daffodil.	☐	☐

The order of the four cases in the partition was random on every trial. The contents of the problems were similar to those in the previous experiment, and they were assigned to the problems in two random ways. Each participant received one of these sets at random, and the eight problems in a different random order.

Results and Discussion

The two groups of participants, who had slightly different instructions, did not differ reliably in either judging truth or estimating probabilities (see the online supplemental materials for the statistics), and so we combined their results for further analysis. Table 4 summarizes the proportion of judgments that each of the cases in a compound's partition could be true. As it shows, the results replicate those of the previous experiment. The complete statistical comparisons are with Table S2 of the online supplemental materials.

Table 5 summarizes the participants' mean estimates of the probabilities of each case in the four sorts of compounds' partitions. Once again, as the model theory predicts, the four sums were subadditive, ranging from 204 to 220. Overall, 66 out of the 81

Table 4
The Mean Proportions of Judgments in Experiment 1b That Each Case in the Partition Could Be True

	Compound assertions			
	Conditionals		Disjunctions	
Truth judgments	Factual	Counterfactual	Factual	Counterfactual
Cases in the partition				
A and C	.98	.96	.99	1.0
Not-A and not-C	.88	.86	.10	.15
Not-A and C	.35	.41	.91	.91
A and not-C	.06	.08	.91	.93

Note. All but one of the proportions (.41) in this table were either significantly greater than or significantly less than the chance value of .5 (Wilcoxon tests, $p < .001$, where Bonferroni corrected alpha is $p < .003$).

participants made subadditive estimates (Binomial test, $p < .001$). The subadditivity of the cases in the partition did not differ significantly depending on the compound: those with disjunctions did not differ reliably from those with conditionals, those with factuals did not differ reliably from those with counterfactuals, and these two variables did not interact reliably (see the online supplemental materials). The subadditivity was likewise smaller than in the previous experiment, though still massive. The reason seems clear: The task focused participants on cases in the partition and thereby reduced the contextual effects of the compounds. Of course, it does not eliminate these effects: Cases in which a compound would be true still elicited much higher probabilities than cases in which it would be false.

A putative alternative explanation is that participants imagined a different background situation for each case in the partition rather than imagining a constant situation. However, we arranged the four judgments vertically beneath each other in a grid to ensure that participants treated them as having a constant background, basing our procedure on an already established one (Goodwin & Johnson-Laird, 2018). The order in which participants received the four cases in the partition was fixed in Experiment 1a and randomized in Experiment 1b, but this factor made little difference. Yet, if participants were entertaining different backgrounds for each comparison, it might be expected to affect their estimates.

Experiment 1c

The previous experiments established robust evidence for the model theory's first two predictions: participants judged that each predicted possibility in a compound's partition could be true and they estimated the corresponding probabilities as subadditive. In Experiment 1a, the participants judged whether a case in the partition and the compound statement could both be true, and estimated the chances that they were both true: Disjunctions were more subadditive than conditionals, and factual conditionals were more subadditive than counterfactual conditionals. In Experiment 1b, in which the compounds provided a context, the participants judged only whether a case in the partition was true and estimated the chances that it was true: Subadditivity was still massive, but its differences from one sort of compound to another were now no longer reliable. Studies of the probabilities of conditionals have sometimes instructed participants that the sum of estimates of cases in the partition should not exceed 100% (e.g., Over et al., 2007). The advantage of this procedure is that it draws attention to the relative probabilities of different cases. Experiment 1c used this procedure.

The model theory predicts that an interaction should emerge reflecting two differences. First, *A and C* is the most salient possibility for factual conditionals—the one case represented explicitly in a mental model (see Table 1), whereas it is merely a counterfactual possibility in the mental models for counterfactual conditionals. Second, *not-A and not-C* has no corresponding explicit mental model for factual conditionals, whereas it is a fact and is represented in a mental model for counterfactual conditionals. Hence, the theory predicts a novel interaction: *A and C* should be more probable for factual than for counterfactual conditionals, whereas *not-A and not-C* should be less probable for factual than for counterfactual conditionals. No sign of such an interaction occurred in the previous studies, perhaps because nothing in them led participants to consider the relative sizes of their estimates, but the task in the present experiment allowed us to test the interaction. It is an instance of the model theory's second prediction, which relates probabilities to possibilities.

Method

Participants. The participants volunteered on the online platform Prolific and were paid £ 1.25 sterling for their participation. We eliminated 26 potential participants for similar reasons to those in the previous studies. There were 72 resulting participants: 49 women, 22 men, and one individual who identified as agender, and their ages ranged from 18 to 76 years with a mean of 35 years.

Design, materials, and procedure. The design, materials, and procedure, were similar to those for Experiment 1b except that the participants were told that the chances for the four statements together must not sum to more than 100, as the following example of the task shows:

> Consider the following *if–then* statement:
>
> If there had been a rose in the vase, there would have been a daffodil.
>
> Now consider each of the four *and* statements on the left below. For each statement, please indicate whether it can be *true*.

Table 5
Mean Estimates of the Probability (From 0 to 100) in Experiment 1b of Each Case in the Partition of the Four Sorts of Compound

	Compound assertions			
	Conditionals		Disjunctions	
Probability estimates	Factual	Counterfactual	Factual	Counterfactual
Cases in the partition				
A and C	80	78	68	67
Not-A and not-C	64	64	29	34
Not-A and C	35	35	56	59
A and not-C	28	27	59	60
Sum	207	204	212	220

Note. The sum of probabilities for each of the four sorts of compound was reliably greater than the correct sum of 100 in the probability calculus (Wilcoxon tests in each case, $z \geq 6.5$, $p < .001$, $r \geq .51$).

Alongside that judgment, please also indicate your estimate of the *chances* that it is true. In the case that you judge that it cannot be true, then plainly the chances that it is true must be 0. The chances for the four statements together *must not* exceed 100.

The *if–then* statement is repeated here:

If there had been a rose in the vase, then there would have been a daffodil.

	Could the statement on the left be TRUE? Write 'Yes' or 'No'	What are the chances that the statement on the left is TRUE? Write a number from 0- 'no chances at all' to 100- 'all chances, each and every one'. Your four answers must sum to 100.
There was a rose and there was a daffodil.	☐	☐
There was a rose and there was NO daffodil.	☐	☐
There was NO rose and there was a daffodil.	☐	☐
There was NO rose and there was NO daffodil.	☐	☐

Results and Discussion

Table 6 summarizes the proportion of judgments that each of the cases in a compound's partition could be true. As it shows, the results corroborated the model theory's first prediction and replicated those of the previous experiments: there was the same trend over the three cases that are true for conditionals, but not for disjunctions. The complete statistical comparisons are with Table S3. in the online supplemental materials.

Table 7 summarizes the participants' mean estimates of the probabilities of each case in the four sorts of compounds' partitions. Of course, the estimates cannot be subadditive, because the participants were told that they must not sum to more than 100. But, the theory predicts the subtle interaction, as the "possibility" tables for the factual and counterfactual conditionals in Table 1 shows, and the results corroborated it. The probability estimates of *A and C* were greater for factual than for counterfactual conditionals, whereas the probability estimates of *not-A and not-C* were smaller for the factual than for the counterfactual conditionals. This interaction was significant as a test of the difference scores revealed (Wilcoxon's test, $z = 3.23$, $p < .001$, $r = .269$), and the two designed comparisons were, too: estimates of *A and C* were greater for factuals than for counterfactuals (Wilcoxon's test, $z = 3.526$, $p < .001$, $r = .294$), whereas estimates of *not-A and not-C* were smaller for factuals than for counterfactuals (Wilcoxon's test, $z = 1.959$, $p < .05$, $r = .163$). No other comparison between factual and counterfactual compounds was reliable (see the online supplemental materials). The results therefore supported an instance of the model theory's second prediction: estimates of the probabilities of cases in the partition reflected the difference between the mental models of factual conditionals and those for counterfactual conditionals. The meaning of factual conditionals parallels counterfactual conditionals, as is clear from their fully explicit models, but the differences in their initial representations in mental models gives rise to the interaction.

Table 6
The Mean Proportions of Judgments in Experiment 1c That Each Case in a Compound's Partition Could Be True

	Compound assertions			
	Conditionals		Disjunctions	
Truth judgments	Factual	Counterfactual	Factual	Counterfactual
Cases in the partition				
A and C	.97	.91	.98	.96
Not-A and not-C	.88	.88	.14	.15
Not-A and C	.42	.43	.94	.90
A and not-C	.11	.13	.94	.92

Note. All but two of the proportions (.42 and .43) in this table were either significantly greater than or significantly less than the chance value of .5 (Wilcoxon tests, $p < .001$, where Bonferroni corrected alpha is $p < .003$).

Table 7
Mean Estimates of the Probability (From 0 to 100) in Experiment 1c of Each Case in the Partition of the Four Sorts of Compound

Probability judgements	Conditionals		Disjunctions	
	Factual	Counterfactual	Factual	Counterfactual
Cases in the partition				
A and C	46	40	34	34
Not-A and not-C	34	38	6	8
Not-A and C	14	14	29	28
A and not-C	5	7	30	30
Sum	99	99	100	99

There were no reliable differences between factual and counterfactual conditionals in judgments of truth, as Table 6 shows. Of course, the *A & C* case is true for both sorts of conditional, even if it has a different epistemic status, and it is represented in the mental models for both, and so the theory predicts no difference. But the *not-A & not-C* case, although true for both sorts of conditional, is represented in the mental models only of the counterfactual and so a difference might have been expected. Previous studies with similar tasks have also failed to find such a difference for neutral content, although it is observed for contents such as causal relations (Thompson & Byrne, 2002). One explanation is that the task explicitly draws participants' attention to the *not-A & not-C* case even for the factual conditional; the difference is more readily observed in tasks that require the participant to generate this case spontaneously, such as inference tasks (e.g., Byrne & Tasso, 1999).

Participants estimated the probability of the *A and C* case for conditionals overall to be high (43 chances out of 100 overall), and the probability of the *A and not-C* case to be low (six chances out of 100 overall), which indicate that asserting a conditional suggests a positive ΔP (the probability of *C* given *A* minus the probability of *C* given *not-A*). However, participants also estimated the probability of the *not-A and not-C* case to be high (36 chances out of 100).

Experiment 2a

The evidence so far has supported the model theory's first two predictions, concerning the truth of cases in the partition and their probabilities. We turn now to an examination of the theory's third principal prediction: judgments of the truth values of compounds should reflect the possibility of counterexamples, but not reflect their likelihood. Hence, a conditional about a particular entity chosen from a set, such as

> If the wine is Italian, then it is red

should be judged to be true for a frequency distribution in which an Italian wine is certain to be red, but it should be judged to be false, or neither true nor false, for frequency distributions in which an Italian wine could be red or white. This latter judgment should occur whether the likelihood of a counterexample, such as a white wine, is high or low. In a previous study, Goodwin (2014) had shown that factual conditionals, both specific and general, tended to be judged as true only in cases of certainty, and not true in any other case regardless of the chances of counterexamples. The aim of the present experiment was to extend Goodwin's findings to counterfactual conditionals, such as

> If the wine had been Italian, then it would have been red.

The experiment accordingly compared three sorts of situation for factual and counterfactual conditionals: those in which the conditional's truth was certain, of a high probability, or of a low probability.

Method

Participants. The participants were 107 volunteers recruited on the online platform Crowdflower (www.figure-eight.com). There were 45 women and 62 men, ranging in age from 18 to 64 years with a mean of 37 years. The participants were recruited from Level 3 on the platform (i.e., those who have completed surveys with a high quality of performance). Another seven participants began the experiment but were eliminated prior to analysis based on screening criteria used in the previous studies. The participants took about 5 min to complete the experiment, and received a nominal payment of US$0.30.

Design and materials. The participants were assigned at random either to factual conditionals ($n = 55$):

> If the wine was Italian, then it was red

or else to counterfactual conditionals ($n = 52$):

> If the wine had been Italian, then it would have been red.

They judged each conditional as true, false, or neither, for three different distributions affecting the conditional probability that a specific wine was red given that it was Italian: certainty, high probability, and low probability. The three distributions were for the frequencies of 100 wines in four categories based on whether a wine was Italian or French and on whether it was red or white. Table 8 presents these three distributions (based on Goodwin, 2014).

Procedure. The materials were presented using SurveyGizmo. The participants were instructed that they would be given three short descriptions to read, and that they had to answer a question about each one. The instructions emphasized that although the three descriptions might seem similar, they were each different, and so they should read each one carefully, paying particular attention to the numbers. A typical trial was as follows:

Table 8
The Three Different Distributions in Experiment 2a for the Probability That an Arbitrary Choice of an Italian Wine Yields One That is Red

The wines		The three distributions		
Country	Color	Certainty	High probability	Low probability
Italian	Red	50	45	30
Italian	White	0	5	20
French	Red	50	20	20
French	White	0	30	30
Total		100	100	100

> The wine list that you are reading lists 100 different wines. Each wine is either from Italy or France, and is either red or white. The following numbers represent the overall frequencies of the different kinds of wine that are available:
>
> | | | |
> |---|---|---|
> | Italian | red | 50 |
> | Italian | white | 0 |
> | French | red | 50 |
> | French | white | 0 |
> | TOTAL | | 100 |

As part of a game with friends, you decide that one of them will choose your wine completely at random from the menu. Imagine that after this selection takes place, someone describes the wine you could have ended up with using the following statement:

> If the wine had been Italian, then it would have been red.

Given the selection of available wines, please indicate whether you think this statement is true, false or neither true nor false:

True ☐

False ☐

Neither true nor false ☐

The participants made their judgment by clicking on one of the three boxes.

Results and Discussion

Table 9 presents the number of participants making the three sorts of judgment for the factual and counterfactual conditionals for three distributions concerning their truth: certainty, high probability, and low probability. The number of participants judging that the conditionals were true did not differ between the factual group and the counterfactual group (Mann–Whitney's $U = 1171$, $z = 1.813$, $p = .07$), and the two groups did not differ reliably for any of the three distributions (see the Fisher's exact tests, $ps > .4$, in the online supplemental materials). But, as the model theory predicts, more participants judged that the certain conditionals were true than that the high and low probability conditionals were true, Cochran's $Q(2) = 73.84$, $p < .001$, Kendal's $w = .345$. For neither the factual conditionals nor the counterfactual conditionals was there any reliable difference for the proportions of true judgments between the high probability and the low probability conditionals (Fisher's exact test, $p > .25$, and $p > .085$, respectively).

In sum, the certain conditionals tended to be judged as true, whereas those of a high or low probability tended not to be (see the

Table 9
The Numbers of Participants in Experiment 2a Judging That Factual and Counterfactual Conditionals Were True, False, or Neither True nor False, in the Certain, High Probability, and Low Probability Situations

Scenarios	True	Neither	False
Factual (n = 55)			
Certain	40	10	5
High probability	17	25	13
Low probability	11	33	11
Counterfactual (n = 52)			
Certain	45	6	1
High probability	20	20	12
Low probability	11	28	13

online supplemental materials, for further statistical comparisons). These results corroborated the model theory's third prediction. They bore out Goodwin's (2014) results with factual conditionals, and showed that they generalize to counterfactual conditionals. The failure to detect a difference in the judgments of high and low probability conditionals is contrary to the probabilistic conditional. People appear to consider a conditional to be true only if its probability is 1. A defense of probabilist views could be that people have different thresholds for characterizing a statement as true depending on the instructions and type of materials (Markovits & Handley, 2005; Oberauer & Wilhelm, 2003; Oberauer, Weidenfeld, & Fischer, 2007). However, conditionals with a high but not certain probability tended not to be judged true reliably more often than false, for factual conditionals: 17 versus 13, $\chi^2(1)$ = .465; and for counterfactual conditionals: 20 versus 12, $\chi^2(1)$ = .157. A counterfactual with a probability of 0 is considered false, and one with a probability of 1 is considered true, and people can use a rich array of terms for counterfactuals whose probability falls between 0 and 1, such as "probably true," "very likely to be false," "almost certainly true" and so on. A high proportion of participants judged that a conditional with high or low probability was neither true nor false rather than false. And, as Table 9 shows, some participants judged that certain conditionals were neither true nor false. Such judgments may have reflected an ambiguity in the instructions, and so we used a different formulation in the next experiment to make them clearer.

According to the model theory, the failure to detect a difference in the judgments of high and low probability conditionals occurs because judgments of truth value do not depend in a systematic way on probabilities less than certainty. Our final experiment examined the model theory's prediction that sensitivity to the probability of counterexamples should occur for judgments on a probabilistic scale.

Experiment 2b

According to the model theory, probabilities tend to be invoked only when the task refers to them. Hence, the present experiment compared two groups of participants who had to make different judgments: one group judged truth or falsity as in the preceding experiment, and the other group made judgments of the probabilities of truth values on a 7-point Likert-type scale, ranging from *certainly true* through *equally probable to be true or false* to *certainly false*. The preceding experiments showed that the results for factual conditionals parallel those for counterfactuals, and so the experiment examined only counterfactual conditionals. It tested them with the same three distributions as those in the preceding experiment: certainty, high probability, and low probability. The model theory's third prediction is that judgments of truth value should not be affected by probabilities: compound assertions are true depending only on the occurrence of instances of the possibilities to which they refer and the nonoccurrence of counterexamples. Hence, they are true only in the distribution in which they are certain. Because the model theory postulates that probabilities tend to be invoked only when the task refers to them, participants who estimate the probability of a conditional's truth values should produce estimates that differ for the three distributions of frequencies.

Method

Participants. The participants were 104 volunteers recruited through the online platform Prolific. There were 28 men, 75 women, and one participant who reported a gender of queer. Their ages ranged from 18 to 65 years with a mean of 33 years. A further three individuals were rejected for the usual reasons before the data analysis. The participants were paid a nominal sum of 50 pence sterling for taking part in the experiment, which took about 5 min.

Design, materials, and procedure. The participants were assigned at random either to the verification group (n = 53), who judged the conditionals as *true*, *false*, *equally possible to be true*, or *it is impossible to say whether the statement is true or false*, or to the probabilistic group (n = 51), who estimated the probability that the conditionals were true or false on the 7-point Likert scale shown below. The participants evaluated each conditional for the three distributions, which were presented in a random order: distributions in which the truth of a conditional was certain, of a high probability, or of a low probability. A typical trial for the verification group was as follows:

> Given the selection of available wines, please indicate whether you think this statement is true, false or equally possible to be true or false. If you think that it is impossible to say whether the statement is true or false, please tick that option instead:
>
> True ☐
>
> False ☐
>
> Equally possible to be true or false ☐
>
> It is impossible to say whether the statement is true or false ☐

As the example illustrates, we disambiguated the *neither true nor false* option used in the previous experiment by providing two options: *equally possible to be true or false* and *it is impossible to say whether the statement is true or false*.

A typical example of a trial for the probabilistic group was:

> Given the selection of available wines, please indicate your level of certainty that this statement is true, false or that it is equally probable to be true or false. If you think that it is impossible to say whether the statement is true or false, please tick that option instead:

−3	−2	−1	0	+1	+2	+3
certainly false	probably false		equally probable to be true or false		probably true	certainly true

It is impossible to say whether the statement is true or false ☐

The three frequency distributions of the wines were the same as those in the previous experiment, but we made one modification to their statement to reinforce a counterfactual interpretation of the conditionals:

> Someone who sees which wine was selected describes the wine you could have ended up with using the following statement:
>
> If the wine had been Italian, then it would have been red.

Hence, the wine you did get was not Italian. The procedure was the same as the previous experiment.

Results and Discussion

Table 10 presents the numbers of participants judging the three sorts of counterfactual conditional in the verification group, and Table 11 presents the numbers of participants making each of the seven judgments on the Likert-type scale in the probabilistic group. As in the previous experiment, more participants judged that the certain conditionals were true than judged that the high and low probability counterfactuals were true, collapsing the results over the two groups, Cochran's $Q(2) = 26.0$, $p < .001$, Kendal's $w = .245$. When the participants judged truth or falsity (see Table 10), there was no reliable difference for judgments of truth between the high and low probability conditionals (Fisher's exact test, $p > .16$).

Most of the responses to the high and low probability conditionals were "equally possible to be true or false" or "impossible to say": 70% for the low probability conditionals (20 + 17 = 37) and 66% for the high probability ones (17 + 18 = 35), as Table 10 shows. These estimates could occur because the participants judge that the conditionals do not fall into the other two available categories of "true" or "false," or because they considered these conditionals to be undecidable.

However, when the participants made estimates of the probability of truth values, the picture changed. As Table 11 shows, conditionals that were certain still elicited judgments of true (+3) more often than highly probable (+2), whereas this pattern switched round for the high probability conditionals (Fisher's exact test, $p < .001$). But, the two distributions for conditionals of high and low probability differed reliably, Kolmogorov–Smirnov two-sample test, $\chi^2(2) = 19.7$, $p < .001$, with more +2 judgments for high probability conditionals, and more +1 judgments for low probability conditionals (Fisher test, $p < .001$).

The results corroborated the model theory's third prediction: Judgments of truth values are based on possibilities rather than frequency distributions, whereas probabilistic estimates of truth values are based on frequency distributions. In both tasks, individuals distinguished between conditionals certain to be true and those that had only a high or low probability of being true. In the verification task, judgments did not discriminate reliably between high and low probabilities. But, in the probabilistic task, judgments were sensitive to frequency distributions. For a distribution in which an Italian wine had a 90% chance of being red, individuals tended to rate a counterfactual conditional:

> If the wine had been Italian, then it would have been red

as highly probable. But, for a distribution in which an Italian wine had only a 60% chance of being red, they tended to rate the counterfactual conditional as only probable. In sum, tasks matter in assessments of truth or falsity: judgments of truth-values reflect only possibilities, whereas estimates of the probabilities of truth values reflect frequency distributions.

Table 10
The Numbers of Participants in Experiment 2b Judging the Truth Values of Counterfactual Conditionals in the Verification Group (n = 53)

Scenarios	True	Equally possible to be true or false	False	Impossible to say
Certain	31	14	4	4
High probability	11	17	7	18
Low probability	6	20	10	17

General Discussion

The model theory postulates that a compound assertion of two clauses, such as a conditional or a disjunction, can refer to certain cases in its partition as factual, possible, counterfactually possible, and impossible. Our five experiments have borne out this account, and corroborated as robust its three principal predictions.

First, participants judged that each of the possible cases in a compound's partition can be true jointly with the compound. The frequency with which such judgments occurred depended on the mental models of the compound. So, judgments for disjunctions did not differ reliably over their three possibilities, whereas judgments of conditionals had a reliable trend based on the accessibility of the cases. For *If A, then C*, the most frequent judgments of truth occurred for the case of *A and C*, which corresponds to an explicit mental model for both factual and counterfactual conditionals. And the case of *not-A and not-C* is more accessible than *not-A and C*, because the former is possible given either a conditional or a biconditional interpretation, whereas the latter is possible only for a biconditional interpretation. Likewise, only the former has an explicit mental model for a counterfactual conditional. The fourth case, *A and not-C*, is impossible, and so it tends to be judged as false. These judgments occurred for joint truth values with the compound (Experiment 1a) and for truth values in the context of the compound (Experiments 1b and 1c). A previous study yielded such results for the joint truth of factual conditionals and cases in the partition (Goodwin & Johnson-Laird, 2018), and so the present experiments extended them to counterfactual conditionals and to disjunctions. Judgments of the truth values of counterfactual compounds ran in parallel to those of factual compounds.

Second, participants' estimates of the joint probabilities of a compound and each case in its partition were grossly subadditive, that is, they summed to over 200% (Experiments 1a and 1b). The model theory predicts this phenomenon. Estimates of the probability of *A*, the probability of C, and the probability of a compound

Table 11
The Numbers of Participants in Experiment 2b Estimating the Probabilities of the Truth Values of Counterfactual Conditionals for Each Value on the Likert-Type Scale, or as "Impossible to Say," in the Probabilistic Group (n = 51)

Scenarios	+3 (*Certainly true*)	+2 (*Probably true*)	+1	0 (*Equally probable to be true or false*)	−1	−2 (*Probably false*)	−3 (*Certainly false*)	*Impossible to say*
Certain	31	4	0	11	1	4	0	0
High probability	1	29	4	8	1	3	4	1
Low probability	1	6	17	13	0	6	4	4
Total	33	39	21	32	2	13	8	5

of *A* and C, tend to be subadditive, because of the intuitive methods on which naive individuals rely, which we described earlier (see Khemlani et al., 2012, 2015). So, estimates of the joint probability of a compound with each case in its partition should elicit four instances of the phenomenon, and lead to the gross subadditivity that occurred in our experiments. As the theory also predicts, disjunctions led to more subadditivity than conditionals. Individuals' intuitive estimates (System 1) of the probability of *A or C, or both*, are a rough average of the probabilities of *A* and *C,* and their deliberative estimates (System 2) tend to add these two probabilities. In contrast, their intuitive estimates of the probability of *If A, then C* nudge the probability of *C* either upward or downward according to their judgment of *A*'s effects, and their deliberative estimates are based on the subset of cases of *A* in which *C* occurs, that is, a correct Bayesian estimate of the conditional probability of *C* given *A*. The difference was robust: the mean estimate of the partition for disjunctions was 18% higher than the mean estimate of the partition for conditionals (Experiment 1a). Likewise, subadditivity was greater for factual than for counterfactual conditionals. Subadditivity, though to a lesser degree, occurred when individuals judged the probabilities of cases in the partition in the context of the compounds (Experiment 1b). But, in this study, there was no reliable difference between disjunctions and conditionals, or between factual and counterfactual conditionals: the task seems to have focused participants on cases in the partition and thereby reduced the effects of the compounds, though cases in which a compound would be true still elicited much higher probabilities than cases in which it would be false.

When individuals were told that their estimates of the probabilities of the four cases in a compounds' partition should sum to no more than 100%, a subtle interaction occurred: the case of *A and C* had higher probability estimates for factual conditionals than for counterfactual conditionals, whereas the case of *not-A and not-C* had lower probability estimates for factual conditionals than for counterfactual conditionals. Of course, *A and C* is a possibility for factual conditionals but a counterfactual possibility for counterfactual conditionals. In contrast, *not-A and not-C* is a fact for counterfactual conditionals but only a possibility for factual conditionals. The interaction did not occur in the first two experiments, perhaps because the need to estimate joint probabilities with compounds led participants to neglect the relative sizes of their estimates.

Third, an arbitrary choice of, say, a bottle of wine, from different frequency distributions can make the truth of a conditional such as:

If the wine had been Italian, then it would have been red

certain, highly likely, or less likely. The model theory predicts that judgments of the truth or falsity of the conditional depend only on the possibility of counterexamples, not their probability, and they did (Experiment 2a). It predicts that compounds refer to possibilities, and so probabilities enter into reasoning only if a task or its contents refers to them. So, one way to invoke them is to ask individuals to rate the probability that an assertion is true. When individuals rated the truth values of conditionals on a seven-point scale, they became sensitive to the difference between a high probability of truth and a low probability of truth. They tended to rate a counterfactual conditional, such as:

If the wine had been Italian, then it would have been red

as +2 (probably true) given that the probability of a red Italian wine was 90%, but only +1 given that the probability of a red Italian wine was only 60% (Experiment 2b).

Overall, the experimental results bore out the model theory's predictions. As it predicts, the results about probabilities showed that the participants made each estimate of cases in a partition independently without realizing that they should sum to no more than 100% (except when so instructed). Other current results also corroborate the theory. From information about what was not the case, for example, "the wine was not red," individuals inferred what was the case, for example, "it was French" in a binary context of Italian or French wine, for factual and counterfactual conditionals. And they inferred what was not the case, for example, "it was not Italian," in a multiple context of Italian or French or Spanish wine (Espino & Byrne, 2018). They selected paraphrases referring to possibilities for factual conditionals and paraphrases referring to counterfactual possibilities for counterfactual conditionals (Quelhas, Rasga, & Johnson-Laird, 2017, 2018). In sum, the two sorts of conditional, factual and counterfactual, differ but do not call for radically different sorts of semantics: Their meanings run in parallel and can be mapped one from the other as in Table 1.

The model theory postulates that compounds refer to conjunctions of possibilities that hold unless there is knowledge to the contrary. Following Zimmermann (2000); Geurts (2005) proposed that disjunctions are treated as conjunctions of alternative possibilities in a semantics based on "possible worlds." The model theory is similar but generalizes the approach to all compounds including conditionals, applies it to counterfactual compounds, bases its semantics on finite alternatives of the sort that hold for everyday probabilities, and stipulates that possibilities hold in default of knowledge to the contrary. Geurts took illusory inferences to hold for conditionals because the correct evaluation of them depended on a truth functional analysis. Our earlier example of an illusion depending on disjunctions, which refer to possibilities, shows that this view is mistaken.

Turning to potential alternative accounts, the suppositional theory of conditionals postulates that conditionals have no truth value in the cases in which their *if*-clauses are false (Evans & Over, 2004). Yet, contrary to this hypothesis, our results show that individuals judge that these cases can be true either jointly with a conditional or given the context of a conditional (Experiments 1a–1c). Perhaps the result could be accommodated in a revision that allowed conditionals to be true or false when their *if*-clause is false, based not only on the de Finetti table but its further specification in the Jeffrey table. But, theories of conditionals based on Adams's (1998) probabilistic conditional (Oaksford & Chater, 2007; Over et al., 2007; Pfeifer & Kleiter, 2010; Oberauer & Wilhelm, 2003) fail to explain several phenomena—the joint truth of a conditional with those cases in its partition in which its *if*-clause is false, the massive subadditivity in estimates of probability, and the paraphrases of conditionals in terms of real possibilities and counterfactual possibilities. The subadditivity in our experiments violates one of the axioms of the probability calculus in its standard formulation, and so it shows that naive human reasoners fail to grasp the foundations of the probabilistic approach to conditionals. Another problem for probabilist approaches is the distinction that individuals make between judgments of truth and estimates of the probability of truth. The former do not reliably reflect differences in less than certain probabilities of conditionals; the latter do (see Experiments 2a and 2b). This distinction should not occur according to a strict interpretation of probabilist theories, unless they invoke pragmatic factors. Not all probabilistic approaches assume that people are always coherent in their probability judgments: They may begin by being less so and their coherence may increase as a result of explicit reasoning in some circumstances (e.g., Evans, Thompson, & Over, 2015).

Another approach outlined earlier is Pearl's (2009, for example) use of Bayesian networks to explain how counterfactual conditionals assert causal relations. In several ways, Pearl's account and the model theory are similar. They concur that probabilities are degrees of belief. They concur that causation is deterministic, not probabilistic, but that ignorance about hidden causal factors forces theorists to introduce probabilities (see, e.g., Chater & Oaksford, 2013). And they also concur that to understand a counterfactual, *If A had happened, then C would have happened*, reasoners simulate what happens by removing *A* from their model of the world, and using knowledge to constrain the simulation (see, e.g., Byrne, 2005; Johnson-Laird et al., 2019). In principle, everything in a Bayesian network can be represented in fully explicit models. The variables in the network and their associated conditional probabilities call for models that have numerical probabilities (see Johnson-Laird et al., 1999). Inferences in the model theory are always defeasible, and so the interventions that underlie causal counterfactuals elicit a process that modifies the original premises, and generates a causal explanation to resolve the conflict between the original premises and the facts of the matter (Johnson-Laird et al., 2015). On the one hand, naive individuals do not know how to compute the probabilities of compound assertions, and so they err in ways that do not occur in a network (Khemlani et al., 2015). On the other hand, they can make kinematic simulations in which loops occur, and loops cannot be represented in Bayes networks (see Khemlani, Mackiewicz, Bucciarelli, & Johnson-Laird, 2013). Likewise, models can represent counterfactuals that are not causal, such as the following (pace Pearl, 2009, Ch. 7):

> If the officer hadn't given the order then the soldiers wouldn't have fired

The soldiers fired because they were under an obligation to obey officers' orders. The order didn't cause them to shoot, because they could have disobeyed it. Bayes networks, of course, make none of the three predictions that our experiments have corroborated. These issues do not matter to Pearl's original goals, which were to show scientists that causation was worth taking seriously, and to establish how to compute the probabilities of causal counterfactuals. However, the issues do count against the direct importation of Bayesian networks into psychological theories of conditionals.

Various future lines of research could be fruitful. One such line is to examine the probabilities of counterfactuals that express causal and other relations. Discoveries about how people make inferences from counterfactuals have generalized from neutral contents about locations, for example, "If Linda had been in Dublin, then Cathy would have been in Galway" (Byrne & Tasso, 1999), to definitional counterfactuals, for example, "If it had been a dog then it would have had four legs" (Thompson & Byrne, 2002), and to causal counterfactuals, for example, "If the water had been heated to 100 degrees, it would have boiled" (Frosch & Byrne, 2012). Likewise, the results of our experiments with neutral contents may extend to the subadditivity of judgments of probability for causal and definitional counterfactuals. However, inferences from deontic contents, for example, "If the nurse had had to clean up blood, she would have had to wear gloves" (Quelhas & Byrne, 2003), and inducements, for example, "If you had tidied your room, I would have given you an ice cream" (Egan & Byrne, 2012) exhibit important differences from other contents. Indeed, Johnson-Laird and Ragni (2019) have argued there can be no probabilist theory of speech acts that create deontic obligations, such as:

> You must tidy your room

You cannot create an obligation, using an assertion such as "The probability that you tidy your room is 1." Hence, an examination of the subadditivity of probability estimates of the truth of deontic counterfactuals could also be informative.

What, then, should readers take away from our research? *If* and *or* refer to real or counterfactual possibilities. Possibilities lie at the roots of the meanings of assertions, and they hold in default of knowledge to the contrary. Facts overturn real possibilities, creating counterfactual possibilities in their place—cases that were once possible but that did not occur. Real possibilities also lie at the roots of the probabilities of assertion, but probabilities introduce differences in the proportions or frequencies with which possibilities occur (Johnson-Laird & Ragni, 2019). Hence, the following inference yields a necessary conclusion:

> It is probable that the Democrats will win the next election.

Therefore, it is possible that the Democrats will win the next election.

But its converse is invalid. Logic and the probability calculus are each incomparable intellectual inventions, and it may be regrettable that they are remote from the roots of human reasoning. In contrast, mental models are heirs to mammalian perception, and they have the advantage of parsimony—they impose a lighter load on working memory than, say, truth tables. But, parsimony is perilous. It leads, as our studies show, to huge overestimates of probabilities, which violate a fundamental axiom of the probability calculus. The meanings of *if* and *or* are not to be found in logic or probability, but in humble possibilities, the forerunners of truth and uncertainty.

References

Adams, E. W. (1970). Subjunctive and indicative conditionals. *Foundations of Language, 6,* 89–94.

Adams, E. W. (1975). *The logic of conditionals: An application of probability to deductive logic*. Dordrecht, the Netherlands: Reidal. http://dx.doi.org/10.1007/978-94-015-7622-2

Adams, E. W. (1998). *A primer of probability logic*. Stanford, CA: Center for the Study of Language and Information.

Barrouillet, P., & Lecas, J. F. (1999). Mental models in conditional reasoning and working memory. *Thinking & Reasoning, 5,* 289–302. http://dx.doi.org/10.1080/135467899393940

Beaver, D. I., & Geurts, B. (2014). Presupposition. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2014 ed.). Retrieved from https://plato.stanford.edu/archives/win2014/entries/presupposition/

Byrne, R. M. J. (1989). Suppressing valid inferences with conditionals. *Cognition, 31,* 61–83. http://dx.doi.org/10.1016/0010-0277(89)90018-8

Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, MA: MIT Press. http://dx.doi.org/10.7551/mitpress/5756.001.0001

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology, 67,* 135–157. http://dx.doi.org/10.1146/annurev-psych-122414-033249

Byrne, R. M. J. (2017). Counterfactual thinking: From logic to morality. *Current Directions in Psychological Science, 26,* 314–322. http://dx.doi.org/10.1177/0963721417695617

Byrne, R. M. J., Espino, O., & Santamaría, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language, 40,* 347–373. http://dx.doi.org/10.1006/jmla.1998.2622

Byrne, R. M. J., & Johnson-Laird, P. N. (2009). "If" and the problems of conditional reasoning. *Trends in Cognitive Sciences, 13,* 282–287. http://dx.doi.org/10.1016/j.tics.2009.04.003

Byrne, R. M. J., & Tasso, A. (1994). Counterfactual reasoning: Inferences from hypothetical conditionals. In A. Ram & K. Eiselt (Eds), *Proceedings of the 16th annual conference of the cognitive science society* (pp. 124–129). Hillsdale, NJ: Erlbaum.

Byrne, R. M. J., & Tasso, A. (1999). Deductive reasoning with factual, possible, and counterfactual conditionals. *Memory & Cognition, 27,* 726–740. http://dx.doi.org/10.3758/BF03211565

Carey, S. (2009). *The origin of concepts*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195367638.001.0001

Chater, N., & Oaksford, M. (2013). Programs as causal models: Speculations on mental programs and mental representation. *Cognitive Science, 37,* 1171–1191. http://dx.doi.org/10.1111/cogs.12062

Costello, F., & Watts, P. (2016). People's conditional probability judgments follow probability theory (plus noise). *Cognitive Psychology, 89,* 106–133. http://dx.doi.org/10.1016/j.cogpsych.2016.06.006

Dehghani, M., Iliev, R., & Kaufmann, S. (2010). Causal explanation and fact mutability in counterfactual reasoning. *Mind & Language, 27,* 55–85. http://dx.doi.org/10.1111/j.1468-0017.2011.01435.x

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2003). Inference suppression and semantic memory retrieval: Every counterexample counts. *Memory & Cognition, 31,* 581–595. http://dx.doi.org/10.3758/BF03196099

Edgington, D. (2014). Estimating conditional chances and evaluating counterfactuals. *Studia Logica, 102,* 691–707. http://dx.doi.org/10.1007/s11225-013-9514-3

Egan, S. M., & Byrne, R. M. J. (2012). Inferences from counterfactual threats and promises. *Experimental Psychology, 59,* 227–235. http://dx.doi.org/10.1027/1618-3169/a000147

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.

Espino, O., & Byrne, R. M. J. (2018). Thinking about the opposite of what is said: Counterfactual conditionals and symbolic or alternate simulations of negation. *Cognitive Science, 42,* 2459–2501. http://dx.doi.org/10.1111/cogs.12677

Espino, O., & Byrne, R. M. J. (2019). *The suppression of inferences from counterfactual conditionals*. Manuscript submitted for publication.

Evans, J. S. B., Thompson, V. A., & Over, D. E. (2015). Uncertain deduction and conditional reasoning. *Frontiers in Psychology, 6,* 398. http://dx.doi.org/10.3389/fpsyg.2015.00398

Evans, J. St. B. T. (1972). Interpretation and matching bias in a reasoning task. *The Quarterly Journal of Experimental Psychology, 24,* 193–199. http://dx.doi.org/10.1080/00335557243000067

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual-processes in reasoning and judgement*. Hove, United Kingdom: Psychology Press. http://dx.doi.org/10.4324/9780203947487

Evans, J. St. B. T. (2012). Questions and challenges for the new psychology of reasoning. *Thinking & Reasoning, 18,* 5–31. http://dx.doi.org/10.1080/13546783.2011.637674

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. http://dx.doi.org/10.1037/0278-7393.29.2.321

Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford, United Kingdom: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198525134.001.0001

Evans, J. St. B. T., Over, D. E., & Handley, S. J. (2005). Supposition, extensionality and conditionals: A critique of Johnson-Laird & Byrne (2002). *Psychological Review, 112,* 1040–1052. http://dx.doi.org/10.1037/0033-295X.112.4.1040

Ferguson, H. J., & Sanford, A. J. (2008). Anomalies in real and counterfactual worlds: An eye-movement investigation. *Journal of Memory and Language, 58,* 609–626. http://dx.doi.org/10.1016/j.jml.2007.06.007

Fillenbaum, S. (1974). Information amplified: Memory for counterfactual conditionals. *Journal of Experimental Psychology, 102,* 44–49. http://dx.doi.org/10.1037/h0035693

Frosch, C. A., & Byrne, R. M. J. (2012). Causal conditionals and counterfactuals. *Acta Psychologica, 141,* 54–66. http://dx.doi.org/10.1016/j.actpsy.2012.07.001

Geiger, S. M., & Oberauer, K. (2007). Reasoning with conditionals: Does every counterexample count? It's frequency that counts. *Memory & Cognition, 35,* 2060–2074. http://dx.doi.org/10.3758/BF03192938

Geurts, B. (2005). Entertaining alternatives: Disjunctions as modals. *Natural Language Semantics, 13,* 383–410. http://dx.doi.org/10.1007/s11050-005-2052-4

Goodwin, G. P. (2014). Is the basic conditional probabilistic? *Journal of Experimental Psychology: General, 143,* 1214–1241. http://dx.doi.org/10.1037/a0034232

Goodwin, G. P., & Johnson-Laird, P. N. (2018). The truth of conditional assertions. *Cognitive Science, 42,* 2502–2533. http://dx.doi.org/10.1111/cogs.12666

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard 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,* 559–569. http://dx.doi.org/10.1037/0278-7393.32.3.559

Hinterecker, T., Knauff, M., & Johnson-Laird, P. N. (2016). Modality, probability, and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 1606–1620. http://dx.doi.org/10.1037/xlm0000255

Jeffrey, R. C. (1981). *Formal logic*. New York, NY: McGraw-Hill.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a theory of meaning, inference and consciousness*. Cambridge, United Kingdom: Cambridge University Press.

Johnson-Laird, P. N. (2006). *How we reason*. New York, NY: Oxford University Press.

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

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109,* 646–678. http://dx.doi.org/10.1037/0033-295X.109.4.646

Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99,* 418–439. http://dx.doi.org/10.1037/0033-295X.99.3.418

Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences, 19,* 201–214. http://dx.doi.org/10.1016/j.tics.2015.02.006

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J.-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106,* 62–88. http://dx.doi.org/10.1037/0033-295X.106.1.62

Johnson-Laird, P. N., Lotstein, M., & Byrne, R. M. J. (2012). The consistency of disjunctive assertions. *Memory & Cognition, 40,* 769–778. http://dx.doi.org/10.3758/s13421-012-0188-2

Johnson-Laird, P. N., & Ragni, M. (2019). Possibilities as the foundation of reasoning. *Cognition*. Advance online publication. http://dx.doi.org/10.1016/j.cognition.2019.04.019

Johnson-Laird, P. N., Ragni, M., Byrne, R. M. J., Khemlani, S., Goodwin, G. P., & Quelhas, A. C. (2019). *The real meaning of If*. Manuscript submitted for publication.

Juhos, C., Quelhas, A. C., & Johnson-Laird, P. N. (2012). Temporal and spatial relations in sentential reasoning. *Cognition, 122,* 393–404. http://dx.doi.org/10.1016/j.cognition.2011.11.007

Khemlani, S. S., Byrne, R. M. J., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model theory of sentential reasoning. *Cognitive Science, 42,* 1–38. http://dx.doi.org/10.1111/cogs.12634

Khemlani, S. S., Lotstein, M., & Johnson-Laird, P. (2012). The probabilities of unique events. *PLoS ONE, 7,* e45975. http://dx.doi.org/10.1371/journal.pone.0045975

Khemlani, S. S., Lotstein, M., & Johnson-Laird, P. N. (2015). Naive probability: Model-based estimates of unique events. *Cognitive Science, 39,* 1216–1258. http://dx.doi.org/10.1111/cogs.12193

Khemlani, S. S., Mackiewicz, R., Bucciarelli, M., & Johnson-Laird, P. N. (2013). Kinematic mental simulations in abduction and deduction. *Proceedings of the National Academy of Sciences, USA, 110,* 16766–16771. http://dx.doi.org/10.1073/pnas.1316275110

Khemlani, S., Orenes, I., & Johnson-Laird, P. N. (2014). The negations of conjunctions, conditionals, and disjunctions. *Acta Psychologica, 151,* 1–7. http://dx.doi.org/10.1016/j.actpsy.2014.05.004

Lagnado, D. A., Gerstenberg, T., & Zultan, R. (2013). Causal responsibility and counterfactuals. *Cognitive Science, 37,* 1036–1073. http://dx.doi.org/10.1111/cogs.12054

Lassiter, D. (2017). *Graded modality: Qualitative and quantitative perspectives*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oso/9780198701347.001.0001

Lewis, D. (1973). *Counterfactuals*. Oxford, United Kingdom: Basil Blackwell.

Liu, I., Lo, K., & Wu, J. (1996). A probabilistic interpretation of 'if–then". *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49,* 828–844.

Lucas, C. G., & Kemp, C. (2015). An improved probabilistic account of counterfactual reasoning. *Psychological Review, 122,* 700–734. http://dx.doi.org/10.1037/a0039655

Markovits, H., Forgues, H. L., & Brunet, M. L. (2010). Conditional reasoning, frequency of counterexamples, and the effect of response modality. *Memory & Cognition, 38,* 485–492. http://dx.doi.org/10.3758/MC.38.4.485

Markovits, H., & Handley, S. (2005). Is inferential reasoning just probabilistic reasoning in disguise? *Memory & Cognition, 33,* 1315–1323. http://dx.doi.org/10.3758/BF03193231

Meder, B., Gerstenberg, T., Hagmayer, Y., & Waldmann, M. R. (2010). Observing and intervening: Rational and heuristic models of causal decision making. *The Open Psychology Journal, 3,* 119–135. http://dx.doi.org/10.2174/1874350101003010119

Meder, B., Hagmayer, Y., & Waldmann, M. R. (2009). The role of learning data in causal reasoning about observations and interventions. *Memory & Cognition, 37,* 249–264. http://dx.doi.org/10.3758/MC.37.3.249

Nickerson, R. (2015). *Conditional Reasoning*. Oxford, United Kingdom: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780190202996.001.0001

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, United Kingdom: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198524496.001.0001

Oaksford, M., & Chater, N. (2017). Causal models and conditional reasoning. In M. Waldmann (Ed.), *The Oxford handbook of causal reasoning*. Oxford, United Kingdom: Oxford University Press.

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 883–899. http://dx.doi.org/10.1037/0278-7393.26.4.883

Oberauer, K., Weidenfeld, A., & Fischer, K. (2007). What makes us believe a conditional? The roles of covariation and causality. *Thinking & Reasoning, 13,* 340–369. http://dx.doi.org/10.1080/13546780601035794

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. http://dx.doi.org/10.1037/0278-7393.29.4.680

Over, D. E. (2009). New paradigm psychology of reasoning. *Thinking & Reasoning, 15,* 431–438. http://dx.doi.org/10.1080/13546780903266188

Over, D. E., & Cruz, N. (2017). Probabilistic accounts of conditional reasoning. In V. A. Thompson & L. J. Ball (Eds), *International handbook of thinking and reasoning* (pp. 434–450). London, United Kingdom: Routledge.

Over, D. E., Hadjichristidis, C., Evans, J. S. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology, 54,* 62–97. http://dx.doi.org/10.1016/j.cogpsych.2006.05.002

Pearl, J. (2009). *Causality* (2nd ed.). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511803161

Pearl, J. (2011). The algorithmization of counterfactuals. *Annals of Mathematics and Artificial Intelligence, 61,* 29–39. http://dx.doi.org/10.1007/s10472-011-9247-9

Pearl, J. (2013). Structural counterfactuals: A brief introduction. *Cognitive Science, 37,* 977–985. http://dx.doi.org/10.1111/cogs.12065

Pfeifer, N., & Kleiter, G. D. (2005). Towards a mental probability logic. *Psychologica Belgica, 45,* 71–99. http://dx.doi.org/10.5334/pb-45-1-71

Pfeifer, N., & Kleiter, G. D. (2010). The conditional in mental probability logic. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 153–173). Oxford, United Kingdom: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199233298.003.0009

Pfeifer, N., & Tulkki, L. (2017). Conditionals, counterfactuals, and rational reasoning: An experimental study on basic principles. *Minds and Machines, 27,* 119–165. http://dx.doi.org/10.1007/s11023-017-9425-6

Quelhas, A. C., & Byrne, R. M. J. (2003). Reasoning with deontic and counterfactual conditionals. *Thinking & Reasoning, 9,* 43–66. http://dx.doi.org/10.1080/13546780244000105

Quelhas, A. C., & Johnson-Laird, P. N. (2017). The modulation of disjunctive assertions. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 70,* 703–717. http://dx.doi.org/10.1080/17470218.2016.1154079

Quelhas, A. C., Rasga, C., & Johnson-Laird, P. N. (2017). A priori true and false conditionals. *Cognitive Science, 41,* 1003–1030. http://dx.doi.org/10.1111/cogs.12479

Quelhas, A. C., Rasga, C., & Johnson-Laird, P. N. (2018). The relation between factual and counterfactual conditionals. *Cognitive Science, 42,* 2205–2228. http://dx.doi.org/10.1111/cogs.12663

Rips, L. J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.

Rips, L. J. (2010). Two causal theories of counterfactual conditionals. *Cognitive Science, 34,* 175–221. http://dx.doi.org/10.1111/j.1551-6709.2009.01080.x

Rips, L. J., & Edwards, B. J. (2013). Inference and explanation in counterfactual reasoning. *Cognitive Science, 37,* 1107–1135. http://dx.doi.org/10.1111/cogs.12024

Santamaría, C., Espino, O., & Byrne, R. M. J. (2005). Counterfactual and semifactual conditionals prime alternative possibilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1149–1154. http://dx.doi.org/10.1037/0278-7393.31.5.1149

Sloman, S. A. (2013). Counterfactuals and causal models: Introduction to the special issue. *Cognitive Science, 37,* 969–976. http://dx.doi.org/10.1111/cogs.12064

Sloman, S. A., & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science, 29,* 5–39. http://dx.doi.org/10.1207/s15516709cog2901_2

Spirtes, P., Glymour, C. N., Scheines, R., Heckerman, D., Meek, C., Cooper, G., & Richardson, T. (2000). *Causation, prediction, and search*. Cambridge, MA: MIT press.

Stalnaker, R. C. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory*. Oxford, United Kingdom: Basil Blackwell. http://dx.doi.org/10.1007/978-94-009-9117-0_2

Thompson, V. A., & Byrne, R. M. J. (2002). Reasoning counterfactually: Making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1154–1170. http://dx.doi.org/10.1037/0278-7393.28.6.1154

Tversky, A., & Kahneman, D. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 292–315. http://dx.doi.org/10.1037/0033-295X.90.4.293

Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural Language Semantics, 8,* 255–290. http://dx.doi.org/10.1023/A:1011255819284

Received November 15, 2018
Revision received July 9, 2019
Accepted July 12, 2019 ■