Cognitive Psychology 122 (2020) 101329

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Conditionals and testimony

Peter J. Collins[a,b,*], Karolina Krzyżanowska[b,c,d,*], Stephan Hartmann[b], Gregory Wheeler[e], Ulrike Hahn[a,b]

[a] *Dept. of Psychological Sciences, Birkbeck, Univ. of London, United Kingdom*
[b] *Munich Center for Mathematical Philosophy, LMU Munich, Germany*
[c] *Institute for Logic, Language and Computation, University of Amsterdam, Netherlands*
[d] *Arché Research Centre, University of St Andrew's, United Kingdom*
[e] *Frankfurt School of Finance & Management, Germany*

ARTICLE INFO

Keywords:
Conditionals
Testimony
Belief updating
Reasoning
Bayesian modeling

ABSTRACT

Conditionals and conditional reasoning have been a long-standing focus of research across a number of disciplines, ranging from psychology through linguistics to philosophy. But almost no work has concerned itself with the question of how hearing or reading a conditional changes our beliefs. Given that we acquire much—perhaps most—of what we believe through the testimony of others, the simple matter of acquiring conditionals via others' assertion of a conditional seems integral to any full understanding of the conditional and conditional reasoning. In this paper we detail a number of basic intuitions about how beliefs might change in response to a conditional being uttered, and show how these are backed by behavioral data. In the remainder of the paper, we then show how these deceptively simple phenomena pose a fundamental challenge to present theoretical accounts of the conditional and conditional reasoning – a challenge which no account presently fully meets.

1. Introduction

Each day we encounter conditional sentences: sentences, most typically, of the form "If P, (then) Q". Conditionals can make assertions: "If people smoke, they have a higher risk of lung cancer." They can give advice or warnings: "If you stand here (there), you'll have a better (worse) view." They can express promises—"If I'm free, I promise I'll come to your party"—and bets—"If Liverpool and Tottenham make the final, I bet you £50 Liverpool will win." They can make claims about worlds that turned out differently from our own—"If Oswald hadn't killed Kennedy, someone else would have"— or that may yet come to pass—"If a sane candidate were to win, we might not be doomed." In short, conditionals are a flexible way to communicate information, or perform speech acts, under uncertainty.

When we receive such conditionals, we must interpret them and we may change our beliefs. These sentences seem natural: we seem to understand them effortlessly and, where appropriate, assimilate their content seamlessly. And yet it remains somewhat mysterious how we change our beliefs when we encounter a conditional. This paper pursues the question of how we change our beliefs in response to conditional assertions. It will focus on *indicative* conditionals: conditional sentences in the indicative mood or, roughly, in which tense has the usual temporal meaning: present means present; past means past; and so on (Declerck & Reed, 2001). Another key type of conditional is the *subjunctive* or *counterfactual* conditional. To see the difference, compare these examples from

* Corresponding authors at: School of Health Sciences, University of Greenwich, Old Royal Naval College, Park Row, Greenwich SE10 9LS, London (P. Collins). Institute for Logic, Language and Computation, University of Amsterdam, Netherlands (K. Krzyżanowska).
E-mail addresses: p.j.collins@greenwich.ac.uk (P.J. Collins), k.h.krzyzanowska@uva.n (K. Krzyżanowska).

https://doi.org/10.1016/j.cogpsych.2020.101329
Received 31 July 2018; Received in revised form 2 June 2020; Accepted 21 June 2020

Declerck and Reed (2001): "I'll punish you if you do something wrong", in which the present tense forms are understood in the usual way, and "I'd punish him if he did something wrong", which can be used of the future despite its past tense form.[1] This latter example is a subjunctive or counterfactual. The tense in such conditionals (often) conveys the falsity of the antecedent. While it is important to understand how people change their beliefs in response to subjunctive conditionals—and much work on subjunctives has been inspired by work on the very type of models we discuss in depth later (Pearl, 2000, 2013)—we defer subjunctives until future work.

There is a further aspect of the mystery that we find important but neglected: the *testimonial* use of conditionals. Conditionals will typically be uttered by a source who (like all of us) is not always right: who is less than perfectly reliable even when well-intentioned. It is the assertion of conditionals by a partially reliable source that provides the basis for recipients to change their beliefs. This paper considers, then, how people update their beliefs when they receive conditional assertions in a testimonial context.

In order to appreciate the overall argument in this paper, it seems helpful to provide an initial broad overview of what our research is attempting to do: in any functional context, it seems essential to one's understanding of what an object is that one takes into consideration what the object does, that is, how one interacts with it. In trying to understand what a tennis racket is, for example, we must observe a game of tennis. Conditionals are no different. There is a long history of trying to understand exactly what conditionals are, in the sense of trying to understand what they mean, but no complete understanding of this can be achieved without engaging with what it is we do with conditionals. For the vast majority of research on conditionals, this has meant examining reasoning with conditionals in the sense of examining the inferences people do and do not draw from a given conditional. But there are other aspects to how we interact with conditionals that have been overlooked: in particular, scant consideration has been given to what changes when we acquire conditionals—specifically how our beliefs change simply on hearing (or reading) a conditional asserted by a testimonial source. As we argue in this paper, consideration of how our beliefs change upon encountering an assertion of a conditional constrains our theories of what conditionals are.

It is by no means unique to the conditional that a theory of what a piece of human language is must engage with the theory of how it is acquired. Historically, it has been integral to debates about the nature of syntax whether or not particular aspects of language might be learnable from the input (Berwick, Pietroski, Yankama, & Chomsky, 2011; Chomsky., 1988; Marcus, 1993). However, it is not just theories of language that are so constrained. Rule-based theories of human general knowledge and common-sense reasoning dominated the early decades of cognitive science (see e.g., Newell & Simon, 1961; Oaksford & Chater, 1991); one of several concerns that brought about their demise was the emergent difficulty of learning suitable rule-based theories (Steels, 1985). In short, theories of what conditionals are must be able to interface broadly with theories of how they are acquired. It is this interface that we explore in the present paper. Specifically, we focus on theories of the conditional that have been influential in the psychology of reasoning, and detail how they fail to interact appropriately with the basic fact that we typically acquire conditionals through testimonial assertion.

For reasons that will become apparent, we will consider how, when people acquire a conditional, they change their judgment of the probability of the conditional's antecendent, the probability of its consequent, and the conditional probability of its consequent given its antecedent. As we will see, some current theories offer straightforward predictions for how our beliefs should change upon hearing such an assertion. These predictions turn out to be wrong. Others offer no direct predictions, but only general 'machinery' for modelling the requisite belief change. That machinery, too, turns out to struggle. This , we argue, suggests that present theories of the meaning of conditionals do not fully square with how we interact with conditionals in our daily lives and that there remains an important gap as a result.

1.1. Background

Although we talk of a mystery about conditionals, the mystery is not total. Conditionals have been widely researched, yielding evidence about what conditionals mean, evidence which bears, of course, on how people change their beliefs in response to a conditional. Much of this evidence argues for a close association between conditionals ("If P, then Q") and the conditional probability of their antecedent given their consequent ($\mathbb{P}(Q|P)$).

One source of evidence is the conditional reasoning task. Participants in these experiments read a set of premises, and either derive their own conclusions or decide whether to endorse conclusions supplied by the experimenter(s). A classic finding is the following, with approximate numbers given in brackets:

Modus Ponens	If P, then Q. P. Therefore Q.	(~100%).
Modus Tollens	If P, then Q. Not Q. Therefore not P.	(~60%).
Affirmation of the Consequent	If P, then Q. Q. Therefore P.	(~50%).
Denial of the Antecedent	If P, then Q. Not P. Therefore not Q.	(~40%).

(see, e.g., Evans, Newstead, & Byrne, 1993).

This pattern indicates that people do not reason according to the rules of classical logic, because classical logic predicts 100% endorsement of *modus ponens* and *modus tollens* and 0% endorsement of Affirmation of the Consequent and Denial of the Antecedent. While a somewhat better fit is achieved by allowing for biconditional interpretations ("if and only if"), a far closer fit is achieved by a probabilistic model which takes conditional reasoning to be Bayesian belief revision with the conditional premise expressed as the conditional probability $\mathbb{P}(Q|P)$ (Oaksford & Chater, 2007). A still closer fit is achieved if we allow violations to the "rigidity"

[1] Note that it is very much up for debate whether the subjunctive mood features at all in "subjunctive" conditionals. We take the view that it does not.

assumption (on the assumption, see Sobel, 2004). That is, learning the minor premise (i.e. the non-conditional premises above) may affect beliefs about the conditional probability. For instance, on hearing “If you turn the key, the car starts” and “The car did not start”, people might lower their judgment of $\mathbb{P}(\text{CarStarts}|\text{KeyTurned})$ (see, e.g., Oaksford & Chater, 2013). In short, the comparative success of the probabilistic models suggests that, however people ultimately represent the conditional,[2] that representation is closely tied to the conditional probability, and that people draw inferences accordingly (on the psychological plausibility of the model assumptions, see Oaksford & Chater, 2013).

Another source of evidence is the judgment task. Participants in these experiments judge the probability of the conditional alongside various other (theoretically relevant) quantities. Studies typically ask whether the probability of the conditional (“If P, (then) Q”) is best predicted by the conditional probability ($\mathbb{P}(Q|P)$), the probability of the conjunction ($\mathbb{P}(P \wedge Q)$), or the probability of the material conditional ($\mathbb{P}(\neg P \vee Q)$). Judgments of the conditional probability correlate well with judgments of the probability of the conditional (Evans, Handley, Neilens, & Over, 2007; Evans, Handley, & Over, 2003; Over, Hadjichristidis, Evans, Handley, & Sloman, 2007b, 2007a; Oberauer & Wilhelm, 2003; Politzer, Over, & Baratgin, 2010). This correlation does not hold universally, however: data suggest that for sixth graders (aged around 12 years’ old) the probability of the conjunction is a better predictor of the probability of the conditional (Barrouillet & Gauffroy, 2015). Studies with adult participants also identify subgroups of participants for whom the probability of the conjunction[3] correlates better with the probability of the conditional (Evans et al., 2003; Oberauer & Wilhelm, 2003; Fugard et al., 2011).[4] Supporting evidence for the role of the conditional probability comes from psycholinguistic studies, which suggest that the conditional probability predicts the reading time of conditionals, and that the conditional probability is available early in processing (Haigh, Stewart, & Connell, 2013).

Studies have shown the association with the conditional probability for a wide range of conditionals. These include indicative conditionals (Evans et al., 2003); causal conditionals (Over et al., 2007b); conditional promises (Ohm & Thompson, 2006); conditional tips, threats, and warnings (Evans, Neilens, Handley, & Over, 2008); and counterfactual conditionals (Over et al., 2007b). Given this breadth of evidence, whatever is learnt from a conditional is likely to include something about the conditional probability.

Rich though these findings are, and suggestive though the link to probabilities is, such studies have not explicitly treated the question of how our beliefs actually change when we hear a conditional. Reasoning research has, instead, focused on reasoning from a set of supplied premises without much regard to how they are acquired. And research on conditionals more generally has focused on establishing a relationship between, on the one hand, the probability, acceptability, and assertability of the conditional and, on the other, the conditional probability without much regard to how someone changes their beliefs when they learn a conditional. These focuses have arguably made the study of the conditional more tractable. But they may also reflect the philosophical origins of reasoning research. Historically, philosophy has modeled reasoning with classical logic, the natural focus of which is the transition from premises to conclusion. Arguably, the psychology of reasoning has inherited this focus, even though many—perhaps most—psychologists of reasoning have adopted new models: witness the probabilistic New Paradigm.[5]

We take this question about belief change to be fundamental to the psychology of reasoning. A comprehensive psychology of reasoning should give an account of reasoning in natural contexts. In such contexts, we contend, reasoners do not typically know all the premises from which they derive their conclusions. They must learn at least some of the premises there and then. Often this learning will occur on the basis of someone’s utterance in a conversation. Indeed, as Oaksford and Chater (2019) observe, “even in the simplest conditional inference, our reasoning depends not on pure principles of logic, but on a combination of background knowledge and conversational principles.” It would surely be problematic for a theory of reasoning if it could not be linked to a model of how the premises are learnt in such contexts (and, as we will see, it is quite plausible to extend probabilistic accounts of reasoning, although the methods we consider produce challenging results).

Against this view, that a full understanding of the psychology of reasoning must take into consideration its interaction with background knowledge and conversational principles, one might set the classic distinction between semantics, pragmatics, and the kind of (putatively logical) reasoning people were traditionally taken to conduct with conditionals. These distinctions might be appealed to in order to support a division of labour, whereby the semantic theory can remain entirely separate from pragmatics, or the reasoning theory distinct from pragmatics. On this alternative view, it is no problem for the semanticist or psychologist of reasoning that their theory omits aspects of belief updating and reasoning; handling those aspects is someone else’s job – say, the pragmaticist’s.

While it is difficult to separate the semantic from the pragmatic (see, e.g., Levinson, 2000; Recanati, 2010; Korta & Perry, 2015), it can certainly be helpful to do so, as the resulting data can help to distinguish between theories of the conditional (for a recent attempt, see Skovgaard-Olsen, Collins, Krzyżanowska, Hahn, & Klauer, 2019). But if we maintain a strict separation, we run the risk of generating component theories that fail to interface with other components in the desired ways. Moreover, as we have seen, it may be impossible to distinguish between reasoning and pragmatics, since conversational principles may inform the most basic inference

[2] We are sympathetic to the view that the conditional is represented at least in part as the conditional probability. But others disagree: see, for example, Goodwin (2014).

[3] There is reason to think that conjunctive responses result from difficulties with mathematical reasoning in these experimental tasks (Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011), and conjunctive responders can learn to respond with conditional probabilities during an experiment (Fugard et al., 2011).

[4] Recent data suggest an important qualification: that, for the conditional probability and the probability of the conditional to correspond well, the antecedent may need to be positively relevant for the consequent (Skovgaard-Olsen, Singmann, & Klauer, 2016b).

[5] There are, of course, influential approaches which use, or are inspired by, logics (see, for instance, Johnson-Laird, Khemlani, & Goodwin, 2015; Stenning & Lambalgen, 2008).

(Oaksford & Chater, 2019). Consequently, in our experiments and modeling, we will not attempt to identify any phenomena as distinctively pragmatic. Some readers might prefer to maintain a separation. But we take the view that strict separation may not only be impossible between the psychology of reasoning and pragmatics, but (even if it were possible) may obscure important constraints on explanatory adequacy.

Similarly, we consider it fundamental that when people learn conditionals—as opposed to, say, probabilistic dependencies from data—they learn from sources that are partially reliable. They learn, that is, from the testimonial use of conditionals. As recipients of information, people are sensitive to sources' reliability from an early age (for a review, see Mills, 2013) into adulthood (for a review, see Briñol & Petty, 2009). It seems essential, then, to understand how this sensitivity to source reliability interacts with the learning of conditionals. This sensitivity may also arise in experiments, however controlled they seem. As Hilton (1995) and Schwarz (1996) have argued, experiments are social and communicative contexts in which a source (the experimenter; like all of us, with partial reliability) communicates information to a recipient (the participant), who need not take that information at face value (for an example, see Hahn, Harris, & Corner, 2016, for discussion of experimenters as sources in studies on climate science communication). Studying the testimonial use of conditionals, then, allows one to consider a crucial aspect of communication, probe deeper into experimental effects, and test the generalizability of experimental findings across levels of trust.

1.2. Some intuitions

How, then, might our beliefs change on hearing a conditional? For some people, this question may have a straightforward answer. Some participants – notably, 6th graders – treat the conditional as a conjunction (Barrouillet & Gauffroy, 2015). For them, learning a conditional presumably amounts to learning the corresponding conjunction. For instance, on hearing "If a car on this lot is a Mercedes, then it's black", they learn "A car on this lot is a Mercedes and it's black." But there appears to be a developmental trajectory away from such conjunctive responding (Barrouillet & Gauffroy, 2015). We will focus here on how *adults'* beliefs change on hearing a conditional, leaving aside the interesting and important question of how children's beliefs change.

Imagine, then, a simple context. You are looking at cars at a large car dealership, and someone tells you "If a car on this lot is a Mercedes, then it's black." Given that there is an association between the conditional and the conditional probability of its antecedent given its consequent, we expect that, when a recipient receives a conditional, they increase this conditional probability. Returning to the car lot example, it seems likely that you would increase your judgment of the probability that a car on this lot is black given that it is a Mercedes. Intuitively, it matters who uttered the conditional: you might well increase the relevant belief if the conditional were uttered by a fellow customer, but you might increase it more if it were uttered by several different people, or if it were uttered by a source with relevant expertise, such as the manager of the car dealership.

This change to the conditional probability is the most obvious intuition, but other changes seem possible. Imagine that you believe today will be a fine, sunny day, but as you prepare to leave your home, someone tells you "If it rains today, then you'll get wet." It seems likely that you would increase your judgment of the probability of rain and, perhaps, of the probability of getting wet. In contrast, imagine that you have applied for a fellowship at a psychology department which you believe you are certain to get. A close friend, who is on the selection panel, tells you "If you get the job, then we will be able to collaborate." Here, it seems possible that you might actually decrease your judgment of the probability of getting the job and, perhaps, of the probability of collaboration. We find it harder to generate intuitions about the probability of the consequent. Our intuitions are weaker and may arise out of coherence: for instance, by *modus ponens*.[6] In the experiments that follow, we treat the probability of the consequent partly for completeness' sake. But the data will also bear on the theories and models we consider. Our choice of theories should be informed by how, on encountering a conditional, people change their judgments of the probability of the antecedent, the probability of the consequent, and the probability of the consequent given the antecedent.[7]

2. The experiments

At present there are only hints in the literature about how people's judgments change on encountering a conditional (see, for instance, Stevenson & Over (2001) and Thompson & Byrne (2002)). For a clear test of the intuitions above and relevant theories, we need new data which isolate testimony.[8] To gather these data, we devised a set of ten simple experiments, which were patterned after a conditional-reasoning task from Stevenson and Over (2001) and standard experimental paradigms for assessing the impact of assertions used in studies on Bayesian argumentation (see, e.g., Hahn & Oaksford, 2007).[9]

As a first step, we used two tasks to probe the impact of assertion on beliefs. We manipulated source characteristics, and as a result

[6] A reviewer identifies a type of conditional where the assertion of a conditional might lead to a change in the conditional probability of the consequent given the antecedent and the probability of the consequent, leaving the probability of the antecedent unchanged. This type is non-interference conditionals, such as "If we triple her salary, Betty will leave the Department" (example due to Douven, 2016a, p.11). These and other conditionals need exploring in future work.

[7] A reviewer argues that a decrease in the probability of the antecedent does not seem intuitively likely in conditionals that express strong inferential relations, such as the following conditional uttered after a winning streak of 9 matches: "If Manchester United win their next match, then they will have won 10 matches in a row.", an intuition that we find plausible.

[8] Stevenson and Over (2001) focus on conditional reasoning, and Thompson and Byrne (2002) compare indicative conditionals with counterfactual conditionals, rather than with a neutral baseline of the kind we describe below.

[9] For the full methodology, see Appendix A.

the extent to which the asserted conditional was likely to be correct. We manipulated this in two different ways. In one set of studies, we manipulated the number of people making the assertion; in the other, the expertise of a single person making the assertion. The designs were between-participants, and across all experiments participants only took part in a single experiment. In both tasks, participants indicated their subjective beliefs in three quantities for a conditional "If P, then Q": these quantities were the probability of the antecedent, $\mathbb{P}(P)$; the probability of the consequent, $\mathbb{P}(Q)$; and the conditional probability, $\mathbb{P}(Q|P)$. Participants responded on a scale from 0 (not at all possible) to 10 (certain).

The first task, the Number-of-Assertions Task, probed participants' response to the mere assertion of a conditional. To illustrate, the following are the different conditions for a single topic, asking about the conditional probability:

Null (No Assertion) Condition: Imagine you are at a large car dealership. What's the probability that a car on this lot is black given that it's a Mercedes?
Single Assertion: Adam is at a large car dealership. He tells you, "If a car on this lot is a Mercedes, then it's black". What is the probability that a car on this lot is black given that it's a Mercedes?
Multiple Assertion: Adam, Barbara, Nick and Sue are at a large car dealership. They tell you, "If a car on this lot is a Mercedes, then it's black?" What's the probability that a car on this lot is black given that it's a Mercedes?

Participants read seven conditionals on real-world topics. For each conditional, they provided three ratings, one for each probability (antecedent, consequent, conditional probability).

The second task probed participants' response to conditionals with source information. To illustrate, the following are the different conditions for a single topic, asking about the probability of the antecedent:

Null Condition: Imagine you are at an infectious-diseases ward. What's the probability that a patient on this ward has malaria?
Inexpert Source: Imagine you are at an infectious-diseases ward. A medical student tells you, "If a patient on this ward has malaria, then they'll make a good recovery." What's the probability that a patient on this ward has malaria?
Expert Source: Imagine you are at an infectious-diseases ward. A professor of medicine tells you, "If a patient on this ward has malaria, then they'll make a good recovery". What's the probability that a patient on this ward has malaria?

Participants read six conditionals on real-world topic[10] in a web survey hosted on Amazon Mechanical Turk (full methods are reported in Appendix A). For each conditional, they provided three ratings, one for each probability (antecedent, consequent, conditional probability). In an initial experiment, we included only the inexpert and expert conditions; from the replication on, we included all three conditions.

In both tasks, we gave the items rather minimal contexts. We intended these contexts to make the items assertable with minimal complexity. We avoided richer contexts to try and isolate the effects of learning a conditional. Richer contexts would require richer representations in any modeling exercise, and would result in less general patterns of belief change.

With these tasks, we explored how people's beliefs change in response to the assertion of a conditional, by manipulating two factors, one per task: Number-of-Assertions (Null, Single, Multiple) and Expertise (Null, Inexpert, Expert). For ease of exposition, we will not present directional predictions here; but, as we will see later, the direction of any changes has considerable theoretical importance.

These initial experiments give us data to assess how beliefs change when participants have no evidence to fix their views on the probability of the antecedent or consequent. The full analysis for this and all other experiments can be found in Appendix A.[11] Let us consider, first, the effect of assertion by single or multiple sources (Number-of-Assertions). The descriptive data are shown in the left-hand panels of Figs. 1–3. An initial data set (Experiment 1) and replication (Experiment 2) produced a consistent pattern of results.[12] Recall that people, here, read conditionals such as "If a car on this lot is a Mercedes, then it's black". Number-of-Assertions reliably increased people's estimates of the conditional probability, $\mathbb{P}(\text{Black}|\text{Mercedes})$. This is borne out by the reliable difference between the Null (context only) and Single conditions ("Adam") and between the Null and Multiple ("Adam, Barbara, Nick, and Sue") conditions. There was however, no difference between the Single and Multiple Conditions. Nor did assertion of the conditional reliably affect the probability of either the antecedent, $\mathbb{P}(\text{Mercedes})$, or the consequent, $\mathbb{P}(\text{Black})$ individually. Table 1, below, reports the fixed effects for Experiments 1 and 2.

We turn, now, to the effect of source expertise (e.g., "medical student" vs. "professor of medicine"). Expertise affected the extent to which the assertion of the conditional increased estimates of the conditional probability. This effect is seen across two data sets, an initial data set (Experiment 3) and replication (Experiment 4). The descriptive data for these studies are shown in the right-hand

[10] We note a potential ambiguity with our items. We intended our items to be about a particular person or object. However, some items could be read as generic statements about any such object or person. Studies on conditionals often use wordings that are more clearly about singular persons or objects (see, for example, Evans et al., 2003). However, Cruz and Oberauer (2014) found evidence that people interpret statements such as our "If a patient on this ward…" conditional as though they are about a random instance from the distribution. Similarly Krzyżanowska, Collins, & Hahn, 2019 found no difference among specific and generic readings of conditionals on a range of variables (both semantic and pragmatic).

[11] Note that the figures show the data aggregated over participants but not over items. They do not, therefore, show the full dependencies in the data, and should be interpreted alongside the parameter estimates in Appendix A. However, the figures and parameter estimates are so close that the figures provide a reasonable and similar summary of the data.

[12] Note that, for both tasks, the confidence intervals were consistent for the original and replication data sets; see Appendix A for more details.

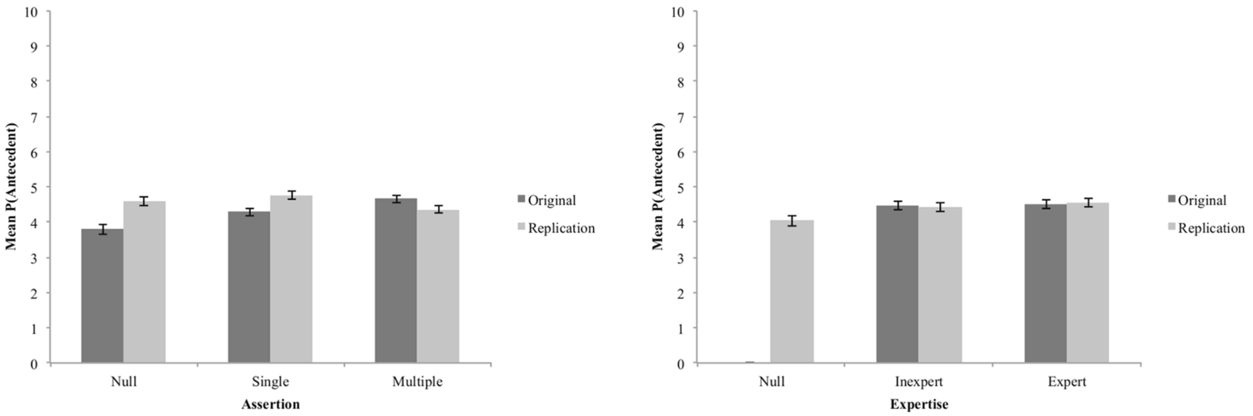

Fig. 1. The Figure shows the descriptive data for ℙ(Antecedent) for the Number-of-Assertions Task in the left-hand panel and the Expertise task in the right-hand panel. The error bars show the standard error.

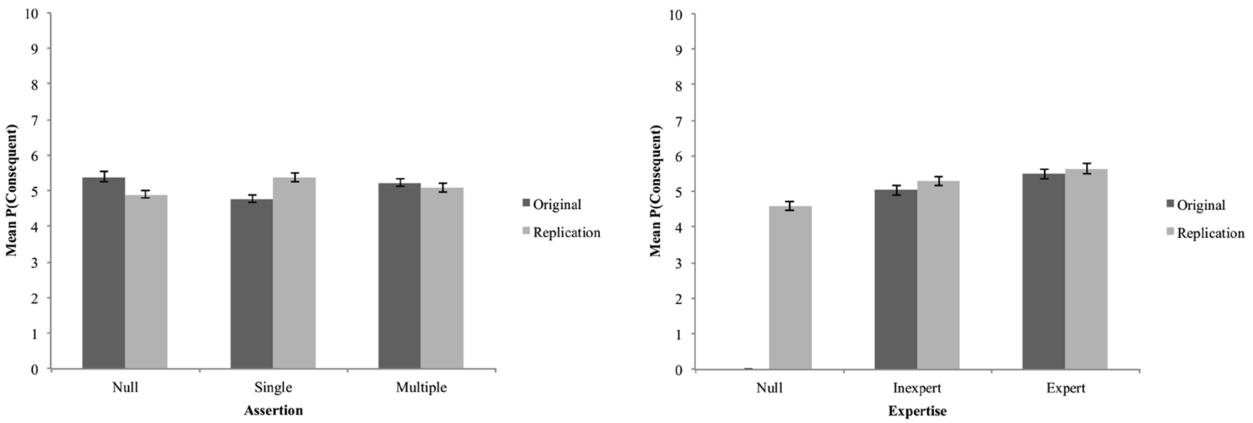

Fig. 2. The Figure shows the descriptive data for ℙ(Consequent) for the Number-of-Assertions Task in the left-hand panel and the Expertise task in the right-hand panel. The error bars show the standard error.

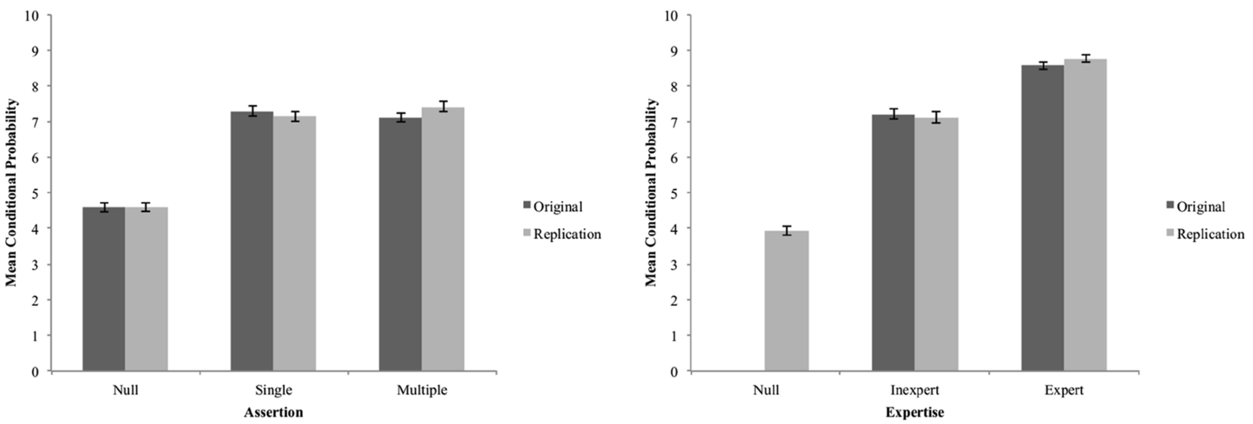

Fig. 3. The Figure shows the data for ℙ(Consequent|Antecedent) for the Number-of-Assertions Task in the left-hand panel and the Expertise task in the right-hand panel.The error bars show the standard error.

Table 1
Test statistics for Experiments 1 and 2; note that CP stands for Conditional Probability.

Experiment	Fixed Effect		Test Statistic	p value
1 – Antecedent	Assertion		$\chi^2(2) = 4.55$	$p = .10$
1 – Consequent	Assertion		$\chi^2(2) = 5.13$	$p = .08$
1 – CP	Assertion		$\chi^2(2) = 11.55$	$p < .001$
		Null – Single	$t(11.04) = 3.63$	$p = .01$
		Null – Multiple	$t(15.87) = 4.38$	$p = .001$
		Single – Multiple	$t(50.52) = -.50$	$p = .87$
2 – Antecedent	Assertion		$\chi^2(2) = 3.03$	$p = .22$
2 – Consequent	Assertion		$\chi^2(2) = 2.41$	$p = .30$
2 – CP	Assertion		$\chi^2(2) = 17.41$	$p < .001$
		Null – Single	$t(46.55) = 6.09$	$p < .001$
		Null – Multiple	$t(24.98) = 5.78$	$p < .001$
		Single – Multiple	$t(103.98) = .77$	$p = .72$

Table 2
Test statistics for Experiments 3 and 4; note that CP stands for Conditional Probability.

Experiment	Fixed Effect		Test Statistic	p value
3 – Antecedent	Expertise		$\chi^2(1) = .01$	$p = .93$
3 – Consequent	Expertise		$\chi^2(1) = 3.28$	$p = .07$
3 – CP	Expertise		$\chi^2(1) = 8.18$	$p < .01$
4 – Antecedent	Expertise		$\chi^2(2) = .77$	$p = .68$
4 — Consequent	Expertise		$\chi^2(2) = 1.34$	$p = .51$
4 – CP	Expertise		$\chi^2(2) = 20.39$	$p < .001$
		Null – Inexpert	$t(10.55) = 6.02$	$p < .001$
		Null – Expert	$t(14.33) = 10.47.$	$p < .001$
		Inexpert – Expert	$t(16.59) = 3.72$	$p = .005$

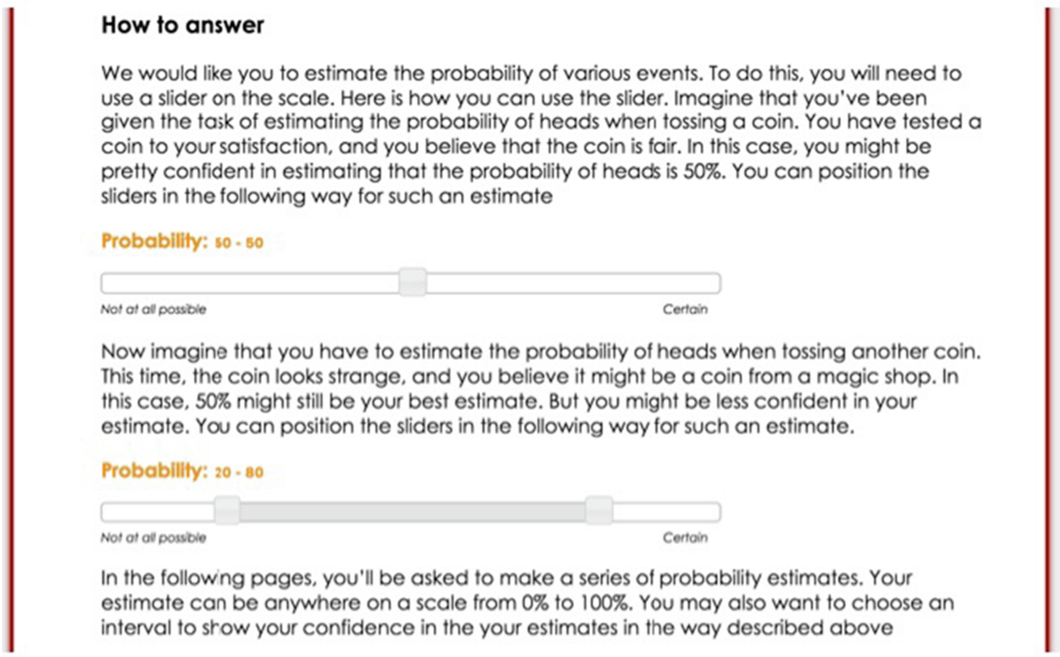

Fig. 4. The Figure shows the instructions and response scale for the intervals studies.

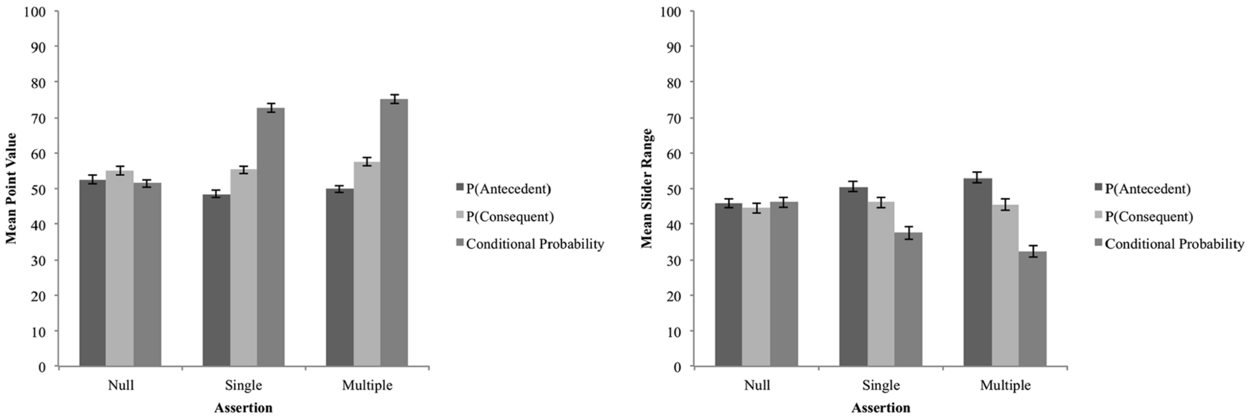

Fig. 5. The Figure shows the effect of Number of Assertions on the three probabilities as measured with sliders. The left-hand panel shows the data as point values. The right-hand panel shows the data as slider ranges. The error bars show the standard error.

Table 3
Test statistics for Experiment 5; note that CP stands for Conditional Probability.

Experiment	Fixed Effect		Test Statistic	*p* value
5 – Antecedent	Point Values		$\chi^2(2) = 1.49$	$p = .47$
	Ranges		$\chi^2(2) = 3.08$	$p = .21$
5 – Consequent	Point Values		$\chi^2(2) = .10$	$p = .95$
	Ranges		$\chi^2(2) = .45$	$p = .80$
5 – CP	Point Values		$\chi^2(2) = 17.76$	$p < .001$
		Null – Single	$t(72.70) = 5.30$	$p < .001$
		Null – Multiple	$t(37.93) = 5.35$	$p < .001$
		Single – Multiple	$t(148.91) = .66$	$p = .79$
	Ranges		$\chi^2(2) = 7.26$	$p = .027$

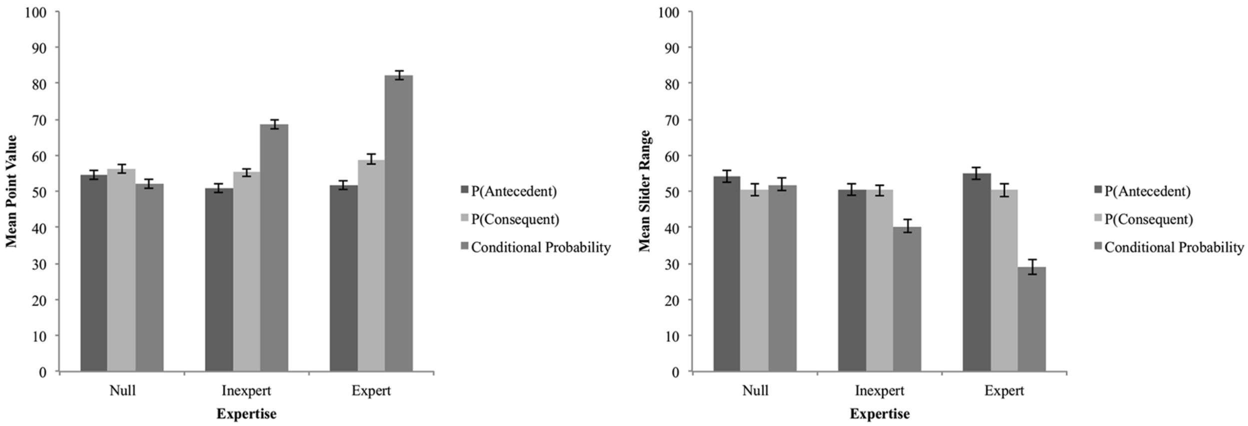

Fig. 6. The Figure shows the effect of Expertise on the three probabilities as measured with sliders. The left-hand panel shows the data as point values. The right-hand panel shows the data as slider ranges. The error bars show the standard error.

panels of Figs. 1–3.[13] Assertion by an inexpert source (a medical student) reliably increased the conditional probability (Experiment 4); so too did assertion by an expert source (a professor of medicine; Experiments 3 an 4); this increase was also reliably larger (Experiment 4). As with number of assertions, neither data set showed reliable effects of source expertise on estimates of the

[13] The initial study included only Inexpert and Expert conditions; the replication included a Null condition to allow more direct comparisons with the Number-of-Assertions study.

Table 4
Test statistics for Experiment 6; note that CP stands for Conditional Probability.

Experiment	Fixed Effect		Test Statistic	p value
6 – Antecedent	Point Values		$\chi^2(2) = 1.88$	$p = .39$
	Ranges		$\chi^2(2) = 1.25$	$p = .53$
6 – Consequent	Point Values		$\chi^2(2) = 1.27$	$p = .53$
	Ranges		$\chi^2(2) = .0008$	$p = 1$
6 – CP	Point Values		$\chi^2(2) = 23.84$	$p < .001$
		Null – Inexpert	$t(75.78) = 4.77$	$p < .001$
		Null – Expert	$t(81.95) = 8.59$	$p < .001$
		Inexpert – Expert	$t(67.75) = 3.79$	$p < .001$
	Ranges		$\chi^2(2) = 16.84$	$p < .001$
		Null – Inexpert	$t(117.56) = -2.05$	$p = .10$
		Null – Expert	$t(152.25) = -4.18$	$p < .001$
		Inexpert – Expert	$t(122.69) = -2.03$	$p = .11$

probability of the antecedent or the consequent. Table 2 reports the fixed effects for Experiments 3 and 4.

There may seem to be a tension between the Number-of-Assertions and Expertise Tasks, in that the effect of Number-of-Assertions seems to plateau, with no reliable difference between Single and Multiple Assertion. But there is reason to believe that this difference is due to the interpretation of the materials in the Multiple condition. The multiple sources can be understood as dependent—e.g. making their observations together—or independent (Wheeler & Scheines, 2013). We pursued this point in another study, reported in Collins (2017). This study included a condition in which participants were told that the sources did not know each other, didn't speak to each other, and gave their testimony independently. In this condition, ratings for the conditional probability were reliably higher than for the single-assertion condition. These additional data suggest that the tension between the Number-of-Assertions and Expertise Tasks is only apparent. We do not discuss the issue of independence further in this paper.

The data from Experiments 1 to 4 suggest that, when people start with intermediate probabilities for the antecedent and consequent, they change only the conditional probability. However, there is another way in which people may change their beliefs on hearing the assertion of a conditional, beyond raising their estimate of the conditional probability. People may not just ascribe a particular degree of (un) certainty to a claim, but also have uncertainty about that degree of uncertainty. In other words, people may be entertaining a range of possible probabilities, and assertion of a conditional may also be affecting those ranges. To probe these effects further, we allowed participants to respond with interval estimates. As above, this was a between-participants design. Participants responded by positioning two end points on a sliding scale between 0% and 100%. Fig. 4 shows the instructions and response scale given to participants.

We replicated the Number-of-Assertions and Expertise Tasks using sliders as the dependent variables. We analyzed the data in two ways: firstly, we averaged the end points of the sliders to replicate the data for the earlier studies (henceforth "point values"); secondly, we analyzed the range of the sliders to ascertain whether the underlying distributions changed (henceforth "slider ranges").[14]

The sliders data broadly replicated the earlier studies. Take, first, assertion by single or multiple sources (Experiment 5). The descriptive data are shown in Fig. 5. When there were multiple sources, there was a tendency for increased precision (decreased ranges) in the conditional probability judgments, but this trend was only marginally significant. The point values data replicated the earlier findings: both single and multiple sources reliably increased judgments of the conditional probability; there was no reliable difference between single and multiple sources. There were no reliable effects on point values or ranges for either the probability of the antecedent or the probability of the consequent. Table 3 reports the fixed effects for this experiment (note the significance level of $p = .025$).

We turn, now, to the the effect of source expertise (Experiment 6). The descriptive data are shown in Fig. 6. Here, there was again a tendency for assertion to increase the precision (decrease the ranges) of conditional probability judgments. But this time this effect was reliable, driven by the difference between Null (context only) and Expert conditions. The point values replicated the earlier data: inexpert and expert sources both reliably increased judgments of the conditional probability, and the effect was reliably larger for expert sources than for inexpert sources. As above, there were no reliable changes to ranges or point values for the probability of the antecedent or the probability of the consequent. Table 4 reports the fixed effects for this experiment.

These intervals data, then, show analogous data patterns to Exp. 1 to 4, which involve point estimates: reliable change only to the conditional probability, both in the mean estimate and (for source expertise) intervals assigned. There were no reliable effects on antecedent or consequent probabilities individually.

But what of the initial intuition that belief change for antecedent and consequent might depend on their prior probabilities? Recall the earlier example of expecting a very dry day and then hearing someone say "If it rains today, then you will get wet." Conceivably this may increase your estimate of the chance of rain. We explored this possibility by manipulating the prior probability of the antecedent by varying the antecedent's fit with a context. Compare, for instance, the following conditionals:

[14] Since these are two analyses of the same data, a multiplicity correction is required. Accordingly we chose a significance level of $p = .025$.

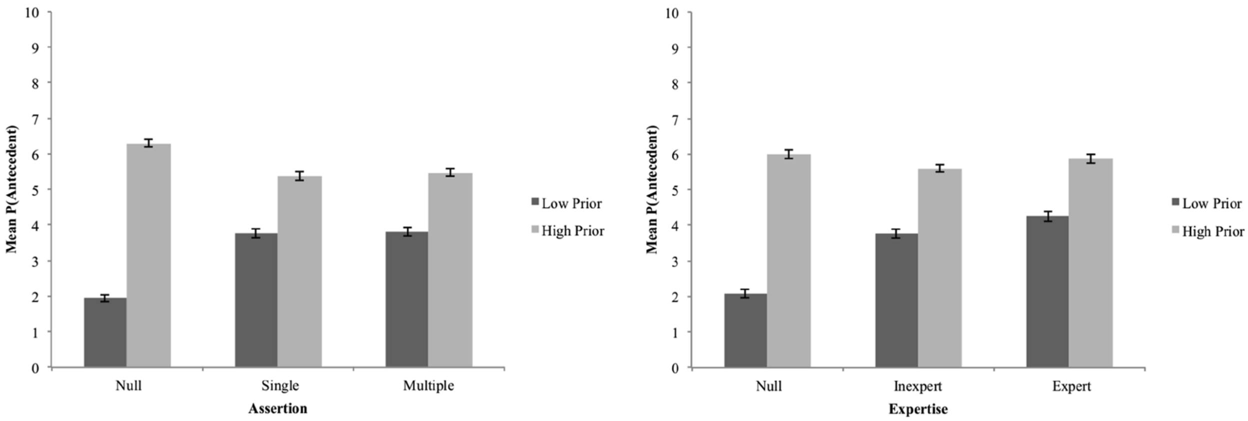

Fig. 7. The Figure shows the effects of Testimony and Prior Probability on ℙ(Antecedent) for the Number-of-Assertions Task in the left-hand panel and the Expertise Task in the right-hand panel. The error bars show the standard error.

Table 5
Test statistics for Experiments 7 and 8; note that CP stands for Conditional Probability.

Experiment	Fixed Effect		Test Statistic	p value
7 – Full	Interaction		$\chi^2(2) = 81.12$	$p < .001$
7 – Low Prior	Assertion		$\chi^2(2) = 18.04$	$p < .001$
		Null – Single	$t(25.39) = 5.99$	$p < .001$
		Null – Multiple	$t(40.71) = 6.13$	$p < .001$
		Single – Multiple	$t(55.24) = .21$	$p = .98$
7 – High Prior	Assertion		$\chi^2(2) = 8.47$	$p = .01$
		Null – Single	$t(14.80) = -2.23$	$p = .10$
		Null – Multiple	$t(32.39) = -3.20$	$p = .008$
		Single – Multiple	$t(25.11) = .37$	$p = .93$
8 – Full	Interaction		$\chi^2(2) = 10.11$	$p = .006$
8 – Low Prior	Expertise		$\chi^2(2) = 13.13$	$p = .001$
		Null – Inexpert	$t(19.29) = 4.48$	$p < .001$
		Null – Expert	$t(19.72) = 5.22$	$p < .001$
		Inexpert – Expert	$t(44.72) = 1.76$	$p = .19$
8 – High Prior	Expertise		$\chi^2(2) = 1.90$	$p = .39$

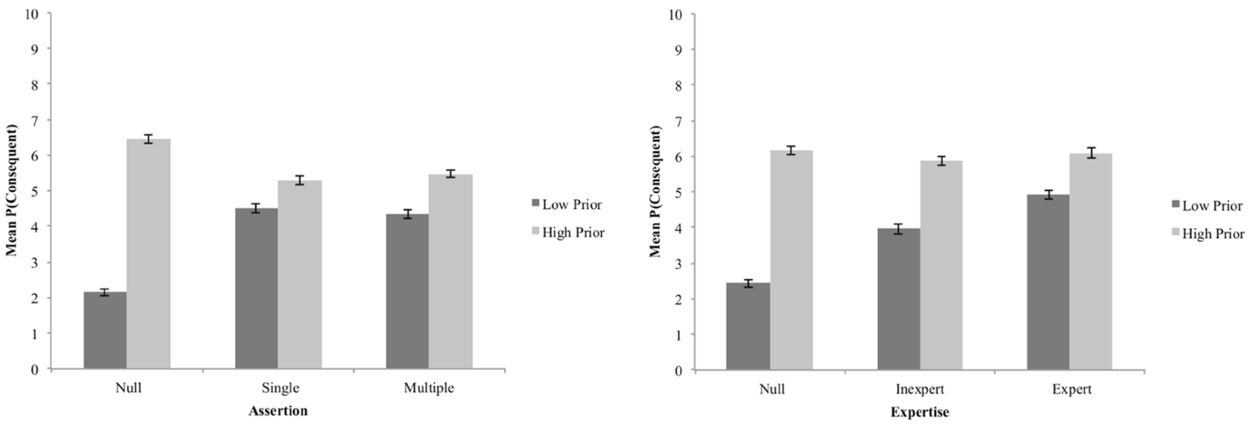

Fig. 8. The Figure shows the effects of Testimony and Prior Probability on ℙ(Consequent) for the Number-of-Assertions Task in the left-hand panel and the Expertise Task in the right-hand panel. The error bars show the standard error.

Table 6
Test statistics for Experiments 9 and 10; note that CP stands for Conditional Probability.

Experiment	Fixed Effect		Test Statistic	p value
9 – Full	Interaction		$\chi^2(2) = 120.16$	$p < .001$
9 – Low Prior	Assertion		$\chi^2(2) = 18.19$	$p < .001$
		Null – Single	$t(27.87) = 7.05$	$p < .001$
		Null – Multiple	$t(15.85) = 5.78$	$p < .001$
		Single – Multiple	$t(41.48) = .10$	$p = .99$
9 – High Prior	Assertion		$\chi^2(2) = 5.48$	$p = .06$
10 – Full	Interaction		$\chi^2(2) = 48.66$	$p < .001$
10 – Low Prior	Expertise		$\chi^2(2) = 20.79$	$p < .001$
		Null – Inexpert	$t(33.39) = 4.33$	$p < .001$
		Null – Expert	$t(35.64) = 7.51$	$p < .001$
		Inexpert – Expert	$t(41.67) = 3.57$	$p = .003$
10 – High Prior	Expertise		$\chi^2(2) = 1.12$	$p = .57$

Table 7
The table summarize the results reported in Section 2 on the effect of learning “If A, C” on participants’ judgments of $\mathbb{P}(C|A)$, $\mathbb{P}(A)$ and $\mathbb{P}(C)$. These results constitute constraints that the data impose on any theory of learning from conditional information that aspires to descriptive accuracy. The asterisk (*) indicates that the increase in probability is only marginally significant.

Prior (no testimony)		Posterior (after learning that “If A, C”)
$\mathbb{P}(C\,|\,A)$	$<$	$\mathbb{P}_{\text{“If A,C”}}(C\,|\,A)$
very low $\mathbb{P}(A)$	$<$	$\mathbb{P}_{\text{“If A,C”}}(A)$
$\mathbb{P}(A)$	$=$	$\mathbb{P}_{\text{“If A,C”}}(A)$
very high $\mathbb{P}(A)$	$\geqslant^*$	$\mathbb{P}_{\text{“If A,C”}}(A)$
very low $\mathbb{P}(C)$	$<$	$\mathbb{P}_{\text{“If A,C”}}(C)$
$\mathbb{P}(C)$	$=$	$\mathbb{P}_{\text{“If A,C”}}(C)$
very high $\mathbb{P}(C)$	$\geqslant^*$	$\mathbb{P}_{\text{“If A,C”}}(C)$

Imagine that you are visiting a Liberal Arts College.
Low Prior: Sue tells you, ‘If Lisa, a student, is majoring in astrophysics, then she’s working late in the library.

High Prior: Sue tells you, ‘If Lisa, a student, is majoring in an arts subject, then she’s working late in the library.’

In this study, then, we manipulated the prior probability and testimony (in one experiment, Number-of-Assertions; in the other, Expertise). Participants were recruited in the same manner as the previous experiments. Participants once again responded on a scale from 0 (not at all possible) to 10 (certain), this time providing a rating only for the probability of the antecedent; and, once again, participants only saw one condition and completed only one task. The key prediction was an interaction between prior probability and testimony: a possible increase when the prior probability is low, and potentially a decrease when the prior probability is high.

The data partially vindicate the intuitions about priors.[15] We will consider, first, the probability of the antecedent (Experiments 7 and 8). The descriptive data are shown in Fig. 7. In both experiments, there were significant interactions. In both experiments, when the prior probability of the antecedent was low, testimony (Number-of-Assertions, Expertise) reliably increased the probability of the antecedent. For Number-of-Assertions (Experiment 7), there were reliable increases between the Null and Single conditions and the Null and Multiple conditions, but not between the Single and Multiple conditions. For Expertise (Experiment 8), there were reliable increases between the Null and Inexpert conditions and between the Null and Expert conditions, but not between the Inexpert and Expert conditions. When the prior probability of the antecedent was high, the tasks differed. Number of assertions reliably decreased the probability of the antecedent, an effect driven by the decrease from Null to Multiple conditions, the other differences not being significant. Expertise, in contrast, did not reliably decrease the probability of the antecedent. Table 5 reports the fixed effects for these experiments (note the significance level of $p = .017$).

We turn, now, to the probability of the consequent. As above, the prior was determined by the fit between context and consequent. The materials were the same as for the task above; we simply swapped the antecedent and consequent.[16] As above, the key prediction was for an interaction between the prior probability of the consequent and testimony (Number-of-Assertions, Expertise).

[15] For all of the priors studies, we selected a significance level of $p = .017$ to correct for multiple comparisons (in addition to the Tukey correction on pairwise comparisons).

[16] There was one exception, where a bigger change was required because simple inversion resulted in a highly unnatural sentence. For details, see Appendix A.

The descriptive data are summarized in Fig. 8.

This study produced similar data for both experiments (Experiments 9 and 10). In both experiments, the interaction was significant. And in both experiments, when the prior probability of the consequent was low, testimony (Number-of-Assertions, Expertise) reliably increased the probability of the consequent. For Number-of-Assertions (Experiment 9), the increase from Null to Single conditions was reliable, as was the increase from Null to Multiple conditions. The increase from Single to Multiple conditions was not reliable. For Expertise (Experiment 10), the increases from Null to Inexpert, from Null to Expert, and Inexpert to Expert were all reliable. For both tasks, when the prior probability of the consequent was high, there was no reliable effect of testimony, in the sense of Number-of-Assertions or Expertise, at the $p = .017$ level. Table 6 reports the fixed effects for Experiments 9 and 10.

2.1. Summary

Across ten experiments we found that participants systematically changed their judgments of the conditional probability, increasing it in response to testimony (Number-of-Assertions, Expertise). Participants were also prepared to make their beliefs more precise, in the sense that they narrowed their interval judgments of the conditional probability (reliably for Expertise, and marginally so for Number-of-Assertions). And they were prepared to change their judgments of the probability of the antecedent and the probability of the consequent depending on the respective priors. For the antecedent, when the priors were low, testimony (both Number-of-Assertions and Expertise) reliably increased the probability; when the priors were high, Number-of-Assertions reliably decreased the probability, but Expertise had no reliable effect. For the consequent, when the priors were low, testimony (both Number-of-Assertions and Expertise) increased the probability; when the priors were high, there was no reliable change.

The effects above seem intuitive and simple. Nevertheless, in the rest of this paper, we will argue that these data pose challenges for both modeling and present theories of the meaning of the conditional. These data are, to the best of our knowledge, novel and, as such, require (further) replication. We do not claim that the findings are general across conditionals. A larger research program is needed to explore patterns of belief change across the rich typology of conditionals, following, for instance, corpus linguistic work such as Declerck and Reed (2001). A program of this kind could usefully cover a wider range of prior probabilities. It should also include a wider range of participants, and could usefully compare adult and developmental groups. However, the data seem an appropriate starting point.

Below, we start by considering how the data sit with the dominant, probabilistic theory of the conditional. We consider two possible extensions, using distance measures and Bayesian belief networks, which have limited success. We then consider how the data challenge other leading theories of the conditional.

3. The data meet the suppositional theory

Any theory whose aspiration is to capture the way people reason with indicative conditionals needs to be compatible with data such as ours on belief change, too (for a summary of the findings, see Table 7). In the Background section we saw considerable evidence for an association between a conditional and the conditional probability of its antecedent given its consequent. These data sit most comfortably with a theory that takes the conditional to be fundamentally probabilistic: the Suppositional Theory. It is natural to think of the Suppositional Theory as the most promising candidate for a theory that can handle these data. However, this section will show that, to handle such data, we will need to extend the machinery of the Suppositional Theory.

What is known as the Suppositional Theory is more accurately a family of probabilistic accounts of conditionals that emphasise the association between the probability of a conditional and the corresponding conditional probability. More specifically, the Suppositional Theories are committed to Stalnaker's Hypothesis (Stalnaker, 1970), which—reflecting its importance in the psychology of reasoning—is widely referred to simply as "the Equation", presented here in a form familiar from the literature on conditionals:

Stalnaker's Hypothesis ("The Equation"): For all probability functions $\mathbb{P}$ such that $\mathbb{P}(A) > 0$, $\mathbb{P}(\text{If}A,\ C) = \mathbb{P}(C|A)$.[17]

As we have explained above, there is considerable evidence that people's interpretation of conditionals complies with the Equation, at least when the conditional's antecedent is positively relevant for its consequent.[18] These data, however, lead to a puzzle, as they seem to clash with an important formal finding, the Triviality Results. Lewis (1976), and other authors after him, undermine the equating of $\mathbb{P}(\text{If}A,\ C)$ and $\mathbb{P}(C|A)$ by showing that it holds only for special, trivial probability functions which cannot be realistic representations of people's degrees of beliefs. To respond to this clash, many researchers choose not to drop "the Equation" and instead drop one of the main assumption of triviality proofs, namely, the idea that conditionals express propositions, that is, that they can be true or false at all (Adams, 1975; Bennett, 2003; Edgington, 1995b). Others choose to follow de Finetti (1937), who analysed conditionals as *conditional events*, satisfying the three-valued truth table, as shown below:

[17] In formulating Stalnaker's Hypothesis this way, we follow a tradition in the literature on conditionals (e.g. Bennett, 2003; Edgington, 1995a; Lewis, 1976; and Stalnaker, 1976). The condition $\mathbb{P}(A) > 0$ can be dropped if one takes conditional probabilities as primitive (e.g. Dubins et al., 1975; Popper, 1959; Rényi, 1955)

[18] Cf. Skovgaard-Olsen, Singmann, and Klauer (2016a)

A	C	If A then C
True	True	True
True	False	False
False/Void	True/Void	Void
False/Void	False/Void	Void

On this account, a conditional is construed as having truth values only when its antecedent is true. The conditional is then true when its consequent is also true, and false when its consequent is false. When the antecedent is false (similarly, when its truth value is undetermined), the conditional does not have a truth value; it is "void," in the same way as a bet on, say, a coin landing heads has no winner if the coin is not thrown at all (de Finetti, 1936/1995, 1974).[19]

The latter approach became predominant in the psychology of reasoning, owing to the considerable evidence from experiments using the truth-table task (e.g. Baratgin, Politzer, & Over, 2013; Over & Baratgin, 2017; Over & Cruz, 2018). In the truth-table task, for a conditional "If A, then C", participants might be asked which A and C states (i.e. true/false) render the conditional true or would (dis-) prove a conditional rule. It is often reported, here, that, when A, the antecedent, is false, people prefer to say that the conditional is neither true nor false but, rather, is irrelevant or void (Over, 2016; Politzer et al., 2010).[20] This result, known as the "defective truth table of the conditional," is argued to be best captured by de Finetti's three-valued system (Baratgin, Politzer, Over, & Takahashi, 2018).

Since de Finetti's third value is to be interpreted as representing a state of ignorance, his logic can be naturally augmented with subjective probability (de Finetti, 1974). The resulting truth table for conditionals, known as the Jeffrey table (e.g. Over & Baratgin, 2017; Over & Cruz, 2018), renders conditionals as true or false, when their antecedents are true, depending on the truth value of their consequents. The false antecedent cases, when the conditional's objective truth value is undetermined or "void," receive then a subjective probability assignment, that is, an agent's degree of belief in a given statement. Consequently, when the antecedent is false or undetermined, the evaluation of a conditional amounts to evaluation of the conditional probability of its consequent given the antecedent.

Since conditionals are often—perhaps even most typically—used when there is uncertainty about the truth of their clauses (see, e.g., Elder & Jaszczolt, 2016), the question of how people estimate the corresponding conditional probabilities became the central question of the "New Paradigm" psychology of reasoning. The psychological process of fixing one's subjective degrees of belief in a conditional has been argued to resemble the Ramsey Test, that is, a procedure for deciding whether or not to accept an indicative conditional suggested by Frank Ramsey in his famous footnote:

> If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q: so that in a sense 'If p, q' and 'If p, $\neg q$' are contradictories. We can say they are fixing their degrees of belief in q given p (Ramsey, 1929/1990, p. 155).

What the above passage suggests are the acceptability conditions for indicative conditionals which can straightforwardly be understood in probabilistic terms: the degree to which we find that "if A, then C" is acceptable, "goes by," or is correlated with, our subjective degree of belief in C given A, that is, the corresponding conditional probability (Adams, 1975).[21]

Usually, unless we have reasons to believe otherwise, we take speakers to assert what they believe to be true, or at least highly likely (see, e.g., Pagin, 2016). In other words, a speaker's assertion that P conveys that the speaker has a high degree of belief—a high subjective probability value—for P being true. Analogously, whenever a speaker asserts a conditional, it follows from the suppositional view on conditionals that they convey that their subjective probability of that conditional's consequent is high under the supposition of its antecedent. It is natural then to understand the learning in response to an assertion of a conditional, that is, a testimony that "If A, then C," assuming that the speaker is sufficiently reliable, as learning that $\mathbb{P}(C|A)$ is high. Thus, the account seems to correctly predict that learning a conditional will result in the increase of the corresponding conditional probability.

However, another consequence of the suppositional account is that we might not be able to directly apply standard probabilistic accounts of learning to the conditional. The Bayesian tradition models the effect of learning a new piece of information, A, on one's degrees of beliefs in terms of *conditionalization*. That is, the posterior probability of any arbitrary proposition C upon learning that A is calculated as equal to the prior conditional probability of C given A:

$$\mathbb{P}_A(C) = \mathbb{P}(C|A) \tag{1}$$

Classical Conditionalization applies when an agent, the learner, accepts A categorically. An agent however may learn A from a not entirely reliable source. In such a case, the agent may change their degrees of belief in A from $\mathbb{P}(A) = .3$ to, for instance, $\mathbb{P}'(A) = .9$. Now, to calculate the posterior probability of any other proposition, the agent cannot simply update by Classical Conditionalization.

[19] See Appendix C for a more detailed discussion of de Finetti's system, and Politzer et al. (2010) on the empirical investigation of the relationship between conditionals and conditional bets.

[20] Cf. Schroyens (2010), who question the strength of this finding. See, also, Douven, Elqayam, Singmann, and van Wijnbergen-Huitink (2018) who found evidence that people are happy to assign truth values to conditionals with false or undetermined antecedents, depending on the presence and strength of an inferential connection between the antecedents and consequents.

[21] Note, however, that while there is significant evidence for "The Equation," the Adams' Thesis, which correlates the degree of *acceptability* of a conditional (since, strictly speaking, one cannot assign probabilities to linguistic constructions that do not express propositions) with the conditional probability of its consequent given the antecedent, has been undermined by the data (Douven & Verbrugge, 2010).

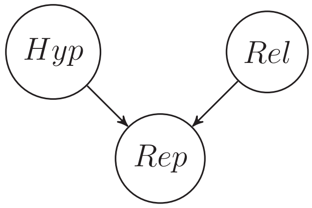

Fig. 9. A Bayesian network for simple (non-conditional) testimony, from Bovens and Hartmann (2003).

Jeffrey (1983) proposed that in such cases the agent should use *Jeffrey Conditionalization* and assign

$$\mathbb{P}'(C) = \mathbb{P}(C|A)\mathbb{P}'(A) + \mathbb{P}(C|\neg A)\mathbb{P}'(\neg A) \tag{2}$$

for the new probability of C.

But it is subject to debate if Classical Conditionalization or Jeffrey Conditionalization can be applied to calculate $\mathbb{P}(A|\text{If A, then C})$ or $\mathbb{P}(C|\text{If A, then C})$, assuming the suppositional interpretation of the conditional. Since the arguments of the probability function need to be propositions, it is not at all straightforward how to calculate the posterior probabilities of A and C, or any other proposition, upon learning that if A, then C, particularly if conditionals are assumed not to express propositions. This is not to say, of course, that no proposals have been put forward. The matter is, however, by no means settled. For instance, Douven and Dietz (2011) show that the Equation (together with another rather compelling assumption) entails $\mathbb{P}(A|\text{If A, then C}) = \mathbb{P}(A)$, which they argue to be a challenge for the Equation itself (see also Douven, 2012). Furthermore, unlike the non-propositional version of the Suppositional Theory, de Finetti's three-valued logic allows conditionals to be embedded as antecedents or consequents of other conditionals.[22] Consequently, as an anonymous reviewer has pointed out, Sanfilippo, Pfeifer, Over, and Gilio (2018) and Sanfilippo, Gilio, Over, and Pfeifer (2020) provide a way to calculate $\mathbb{P}(A|\text{If A, then C})$, though this method is yet to be extended to belief updating.[23]

3.1. Summary

We have seen that the Suppositional Theory lacks the machinery to fully account for how people change their beliefs when they encounter a conditional. Aspects of the theory also mean that it cannot rely on traditional models of belief change in the Bayesian tradition. The question arises, then, of whether we can extend the theory in such a way that we can better capture our data on updating on conditionals without having to forsake a large body of evidence. We turn to other ways to model belief updating. We consider two approaches. The first, the *distance-based approach to Bayesianism*, has emerged as the standard method to model updating on conditionals in formal epistemology (Eva et al., 2019). The second, Bayesian belief networks, adapts a method that does precisely what we want but for assertions of simple propositions.

4. Modeling the data

4.1. Distanced-based Bayesianism

In light of the previous sections, our goal is to model our data while preserving a link with the probabilistic approach to conditionals. A promising way to reach this goal is to model the learning process in the *distance-based approach to Bayesianism*.[24] On this approach, we can understand updating on conditionals as follows: We assume that an agent has beliefs about some set of propositions. For a simple example, let these propositions be A and C, and let the beliefs about these propositions be represented by a prior probability distribution $\mathbb{P}$ over them. The agent now learns new information in the form of a conditional "If A, then C". We ask: How should she change her beliefs in light of this new information? Let us assume, for now, that the agent sets the new conditional probability, $\mathbb{P}'(C|A)$, to 1. The agent now moves to a new set of beliefs about A and C, represented by a new probability distribution $\mathbb{P}'$, and does so by taking the conditional probability assignment $\mathbb{P}'(C|A) = 1$ as a *constraint* on $\mathbb{P}'$. This constraint does not fix the full new probability distribution $\mathbb{P}'$ uniquely. To find $\mathbb{P}'$ and to make sure that her new beliefs are coherent, she then minimizes some appropriate distance measure between $\mathbb{P}'$ and $\mathbb{P}$. The agent is, in a sense, conservative: given the constraint, she changes her other beliefs as little as possible.

To express this account formally, we need some way of measuring the distance between posterior and prior distributions over the

[22] Though see Douven (2016) who showed that, given de Finetti's table, iterated conditionals are equivalent to simple conditionals that are unacceptable in natural language.

[23] For other proposals of update rules, typically designed to accommodate very specific desiderata related to various specific situations, see, e.g., Kern-Isberner (2001), van Fraassen (1981), Bradley (2005), Douven (2012), Douven and Dietz (2011), Douven and Romeijn (2011), Eva, Hartmann, and Rad (2019), Harper (1976), and Gärdenfors (1988).

[24] Following Csiszár (1967), the distance-based approach to Bayesianism was introduced in Diaconis and Zabell (1982). It has been developed and applied to conditionals in Eva and Hartmann (2018), Eva et al. (2019), Sprenger and Hartmann (2019) and Stern and Hartmann (2018). Eva, Hartmann, & Singmann, 2019 use the approach to explain the MP–MT asymmetry mentioned in the Background section.

same algebra of propositions. There are several ways to do so. It turns out that all so-called (statistical) divergencies[25] in the class of *f*-divergences[26] (Csiszár, 1967) imply Jeffrey Conditionalization (see Eq. (2)) if the agent entertains beliefs about two propositions and shifts, as a result of the learning, the probability of one of them to a higher or lower value (see Diaconis & Zabell (1982) and Eva et al. (2019)). If the probability of one of them shifts to 1, then one obtains Classical Conditionalization which, in turn, is a limiting case of Jeffrey Conditionalization. As there are many strong arguments in favor of Classical Conditionalization and Jeffrey Conditionalization (see, e.g., Schwan & Stern (2007) and Jeffrey (1983)) we consider the fact that both follow from minimizing an *f*-divergence (taking a certain constriant into account) to support our focus on this class of divergencies.

Formally, then, an agent who learns the conditional "If A, then C" minimizes an *f*-divergence between $\mathbb{P}'$ and $\mathbb{P}$, taking (in the simplest way of deriving a constraint from the learned conditional) $\mathbb{P}'(C|A) = 1$ as a constraint on $\mathbb{P}'$. Note that this approach is intrinsically probabilistic and, in that sense, is a natural way of extending the probabilistic approach discussed in the previous sections to the learning of conditionals. It also has the advantage that it gives the intuitively correct results for a number of challenging examples put forward by Douven such as the Ski Trip Example (Dietz & Douven, 2010) and the Sundowners Example (Douven & Romeijn, 2011). One can also show that minimizing an *f*-divergence between $\mathbb{P}'$ and $\mathbb{P}$ is equivalent to conditionalizing on the material conditional provided that the corresponding conditional is strict, i.e. if one takes $\mathbb{P}'(C|A) = 1$ as a constraint on $\mathbb{P}'$. For details and proofs, see Eva et al. (2019) as well as the discussion below.

What predictions does the distance-based approach make for updating on conditionals? On the simplest version of this account, the agent fixes the new conditional probability to 1 and finds the full new distribution $\mathbb{P}'$ by minimizing an *f*-divergence between $\mathbb{P}'$ and $\mathbb{P}$. One then obtains that she (i) decreases the judgment of the probability of the antecedent and (ii) increases the judgment of the probability of the consequent (Eva et al., 2019). These qualitative results hold for all prior probability distributions, but the amount of change depends on the specific prior distribution. This, of course, is not at all the pattern we saw in our empirical data, where we saw that probabilities for the antecedent and consequent increased when their respective priors were low, and did not change when they were moderate.

But there is a fundamental problem with the simplifying assumption that we made above: that the agent sets the conditional probability to 1. In none of the experiments did participants, on aggregate, set the conditional probability to 1. But we do not need to make this assumption; an agent raising the conditional probability to a value smaller than 1 can also be modeled in the distance-based approach. In this case, the probability of the antecedent again *always decreases* and the probability of the consequent can decrease, increase or stay the same, at least for the most common *f*-divergencies such as the Kullback–Leibler diveregence, the Hellinger distance and the χ^2-divergence. By contrast, for the inverse Kullback–Leibler divergence, which is also an *f*-divergence, the probability of the antecedent stays the same (For further details, see again Eva et al. (2019)). More success is to be had with models that explicitly incorporate the testimonial context.

4.2. Bayesian belief networks

In this section, we introduce a range of Bayesian Belief Networks which are intended to explicitly model the testimonial context. The details of these models are inevitably somewhat technical. But we will start by introducing a simpler model that can be seen as the basis for the remaining models. For full technical details, we refer readers to the footnotes and Appendix B. For readers less familiar with this form of modelling, we will emphasize the outcomes of the models in brief summary sections so that technicalities can be skipped.

To introduce these models, we first imagine a situation in which a speaker makes a *non-conditional* claim about the world. Imagine that a high-school history teacher makes the claim that Gutenberg invented the printing press. How should we evaluate this claim? Relevant factors are (1) whether we consider a high-school history teacher a reliable source of information and (2) how plausible we find the claim. We can use subjective probabilities to represent both factors: for (1), the prior probability, or base rate, of the particular source telling the truth; for (2), the prior probability of the claim. Being subjective, these probabilities can differ between recipients of the claim: some might trust high-school history teachers more than others; some, perhaps who have read more Chinese history, may have a lower prior probability than others for Gutenberg being the inventor of the printing press.

This simple situation is captured by a model from Bovens and Hartmann (2003) depicted in Fig. 9. This model is a simple Bayesian Network: the nodes represent binary propositional variables, and the arcs represent probabilistic dependencies among the variables. The variable *Hyp* stands for the hypothesis or claim in question and has the values Hyp: "The hypothesis in question is true" and ¬Hyp: "The hypothesis in question in false". For the example above, the "hypothesis" is that Gutenberg invented the printing press. The variable *Rep* stands for the report of the claim and has the values Rep: "The information source reports that the hypothesis in

[25] Statistical divergencies are not strictly speaking distance measures: they may not be symmetrical and they may violate the triangle inequality.

[26] *f*-divergences are defined as follows: Let $S_1, \ldots, S_n$ be the possible values of a random variable S over which a prior probability distributions $\mathbb{P}$ and a posterior probability distribution $\mathbb{P}'$ are defined. The *f*-divergence between $\mathbb{P}'$ and $\mathbb{P}$ is then given by

$$D_f\left(\mathbb{P}'||\mathbb{P}\right) := \sum_{i=1}^{n} \mathbb{P}(S_i) f\left(\frac{\mathbb{P}'(S_i)}{\mathbb{P}(S_i)}\right),$$

where f is a convex function which satisfies $f(1) = 0$. Examples of *f*-divergences are the Kullback–Leibler divergence ($f(x) = x\log x$), which is also widely used in information theory, the inverse Kullback–Leibler divergence ($f(x) = -\log x$), the Hellinger distance ($f(x) = (1 - \sqrt{x})^2$), and the χ^2-divergence ($f(x) = (x - 1)^2$).

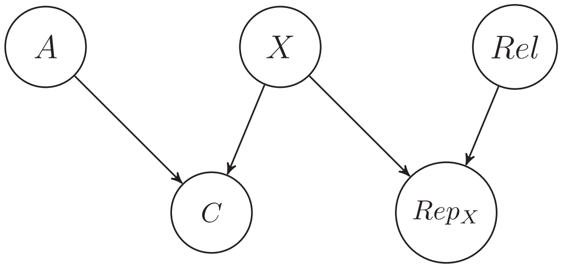

Fig. 10. The Bayesian Network for the Baseline Model.

question is true" and ¬Rep: "The information source reports that the hypothesis in question is false". Lastly, the variable *Rel* stands for the reliability of the source and has the values Rel: "The information source is reliable" and ¬Rel: "The information source is not reliable". For the example above, the information source is the high-school history teacher.

Both the beliefs and reliability are represented by a distribution over the binary states of the respective variables. On this simple model, a reliable source will report that the hypothesis in question is true (i.e. Rep) if the hypothesis is true (Hyp) and will report that the hypothesis in question is false (i.e. ¬Rep) if the hypothesis is false. An unreliable source, by contrast, will 'randomize' and report Rep independently of the truth or falsity of the hypothesis in question—essentially flipping a (possibly biased) coin. This can be captured in the following prior probability assignments:

$$\mathbb{P}(\text{Rep}|\text{Hyp}, \text{Rel}) = 1 \quad , \mathbb{P}(\text{Rep}|\neg\text{Hyp}, \text{Rel}) = 0$$
$$\mathbb{P}(\text{Rep}|\text{Hyp}, \neg\text{Rel}) = \mu \quad , \quad \mathbb{P}(\text{Rep}|\neg\text{Hyp}, \neg\text{Rel}) = \mu$$

Here μ is the so-called randomization parameter. To complete the Bayesian Network, we also have to assign prior probabilities to the root nodes *Hyp* and *Rel*:

$$\mathbb{P}(\text{Hyp}) = h, \quad \mathbb{P}(\text{Rel}) = r$$

Of course, real-world sources are unlikely to simply pick an answer at random, but the model is intended to provide a minimal model of an agent who provides us with testimony, but whose reliability is not exactly known. The randomization parameter simply allows one to capture the fact that to the extent that this agent is unreliable, their report is determined independently of the true state of the world. In other words, the model seeks to capture the essential situation of a recipient who knows little about her testimonial source. More detailed, psychologically plausible extensions to this model are possible where suitable evidence about expertise or bias exist (see e.g., Harris, Hahn, Madsen, & Hsu, 2016). Both these more elaborate models and the basic version presented here share the same fundamental, qualitative characteristics: evidence reports affect both beliefs about the hypothesis or claim in question and beliefs about the reliability of the source. Expected evidence will increase both belief in the hypothesis and in the reliability of the source; unexpected evidence will lead to an increase in belief in the claim, but will decrease belief in the reliability of the source. These basic dynamics of testimonial evidence have been confirmed in behavioral studies (Collins, Hahn, von Gerber, & Olsson, 2018). They have also formed the basis of experimental work in the studies of Jarvstad and Hahn (2011), Hahn, Harris, and Corner (2009), and Harris et al. (2016).

Although it might be tempting to assign the conditional to node *Hyp*, this strategy is problematic on both conceptual and empirical grounds. The strategy is conceptually problematic because it requires us to ignore the fact that conditionals refer, in some way, to a relationship between the antecedent and consequent propositions.[27] The strategy is empirically problematic because it would not allow us to capture the separate belief change to antecedent, consequent and conditional probability.

Instead, we propose two models that can be seen as extensions of the above model of testimony. These models allow us to study the learning of an indicative conditional "If A, then C" from a partially reliable information source. To do so, we introduce the following five binary propositional variables.[28]

1. The variable *A* has the values A: "The antecedent occurs" and ¬A: "The antecedent does not occur".
2. The variable *C* has the values C: "The consequent occurs" and ¬C: "The consequent does not occur".
3. The variable *X* has the values X: "The indicative conditional 'If A, then C' holds" and ¬X: "The indicative conditional 'If A, then C' does not hold". (Extending the algebra and introducing the conditional as an additional variable is a trick that goes back to Garber's solution to the problem of old evidence.[29])
4. The variable Rep_X has the values Rep_X: "The information source reports X" and $\neg\text{Rep}_\text{X}$: "The information source does not report X."
5. The variable *Rel* has the values Rel: "The information source is reliable" and ¬Rel: "The information source is not reliable."

[27] We do not have anything specific in mind by using the term "relationship" here. The term could merely refer to the conditional probability of one proposition given another. We are not committing to conditionals conveying by default an inferential or causal relationship.

[28] We use italics to denote the variables and roman script for the instantiations of the variables.

[29] See Garber (1983) as well as Hartmann & Fitelson, 2015; Sprenger & Hartmann, 2019; Eva and Hartmann, 2019 for recent developments of the account.

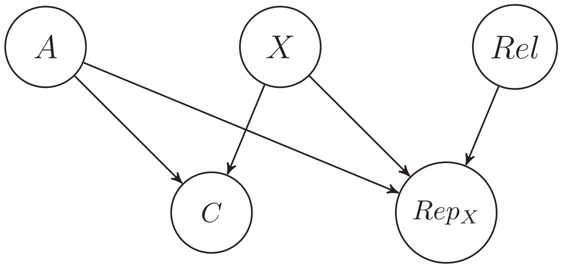

Fig. 11. The Bayesian Network for the Modified Model A.

To proceed, we represent the probabilistic relations between these variables in a Bayesian Network. We propose five different models for this: We begin with the Baseline Model and show that this model does not account for everything we would like to explain. A series of modified models are then shown to go some way towards doing the job. However, none can fully capture our basic data. We discuss these models in turn.

4.3. The Baseline Model

This model assumes the following (conditional) independence relations:

- $A \perp\!\!\!\perp X, Rel$
- $C \perp\!\!\!\perp Rep_X | X$

These (conditional) independence relations are represented in the Bayesian Network in Fig. 10. To complete the network, we specify the probability distribution as follows. First, we specify the prior probabilities of the "root nodes" A, Rel, and X:

$$\mathbb{P}(\mathrm{A}) = a, \quad \mathbb{P}(\mathrm{Rel}) = r, \quad \mathbb{P}(\mathrm{X}) = x \tag{3}$$

Here we assume that $a, r, x \in (0, 1)$.

Next, we specify the probabilities of C, given its parents.

$$\mathbb{P}(\mathrm{C}|\mathrm{A}, \mathrm{X}) = 1, \quad \mathbb{P}(\mathrm{C}|\mathrm{A}, \neg \mathrm{X}) = \alpha \tag{4}$$

$$\mathbb{P}(\mathrm{C}| \neg \mathrm{A}, \mathrm{X}) = \beta, \quad \mathbb{P}(\mathrm{C}| \neg \mathrm{A}, \neg \mathrm{X}) = \gamma$$

Here we assume that the conditional statement X respects *modus ponens*, i.e. that $\mathbb{P}(\mathrm{C}|\mathrm{A}, \mathrm{X}) = 1$. All other parameters (i.e. α, β and γ) are not further specified.

Finally, we specify the probabilities of $\mathrm{Rep_X}$, given the values of its parents.

$$\mathbb{P}(\mathrm{Rep_X}|\mathrm{X}, \mathrm{Rel}) = 1 \quad , \quad \mathbb{P}(\mathrm{Rep_X}| \neg \mathrm{X}, \mathrm{Rel}) = 0 \tag{5}$$

$$\mathbb{P}(\mathrm{Rep_X}|\mathrm{X}, \neg \mathrm{Rel}) = \mu \quad , \quad \mathbb{P}(\mathrm{Rep_X}| \neg \mathrm{X}, \neg \mathrm{Rel}) = \mu$$

To model the report of a partially reliable information source, we assume that the source is reliable with probability *r*. If the source is reliable, then it is a truth teller. If the source is not reliable, then it randomizes with probability μ as described above.

At this point, the reader will likely wonder about the interpretation of node *X*. As we just noted, it is problematic to assign the conditional to node HYP directly, not least because it is problematic to treat the conditional as a proposition (see the earlier discussion of the Triviality Results). Instead we assume simply that when the propositional variable "The indicative conditional 'If A, then C' holds" is true, then C follows from A. This is reflected in the conditional probability table defining the relationship between A, C and X. In other words, we assume simply that the conditional, whatever it may be, respects *modus ponens*. [30]This is relatively non-committal about what the conditional itself *means*. We assume only that when natural language convention licenses one to assert 'if A, then C' then it is acceptable to infer C from A. We do not seek to spell out what those linguistic conventions actually are. Hence we intend these models to be light on theoretical commitments concerning the content and structure of the conditional, focussing instead on an effect of what one believes when one takes the conditional to hold: namely that A entails C.[31]

How does the baseline model bear on the experimental data? We are now in the position to calculate how the probabilities of A (antecedent) and C (consequent) change once we receive the report $\mathrm{Rep_X}$. We will also calculate how the conditional probability

[30] As a reviewer points out, this assumption might be incompatible with some versions of inferentialism, on which *modus ponens* is not valid (e.g. Krzyżanowska, Wenmackers, & Douven, 2014), though see, e.g., Skovgaard-Olsen et al. (2016a) for an alternative approach to inferentiailsm. See also Mirabile and Douven (in press) for an empirical investigation and a helpful discussion of inferential conditionals and *modus ponens*.

[31] We note, however, that there remains a sense in which node *X* refers to a link between nodes *A* and *C* of the same network, raising the question of whether we are conflating object- and meta-languages in the same representation. In Appendix C, we argue that this can nevertheless be given a coherent, probabilistic interpretation.

$P(\text{C}|\text{A})$ changes after learning Rep_X. To do so, we define

$$\begin{aligned} \Delta_A &:= \mathbb{P}(\text{A}|\text{Rep}_\text{X}) - \mathbb{P}(\text{A}) \\ \Delta_B &:= \mathbb{P}(\text{C}|\text{Rep}_\text{X}) - \mathbb{P}(\text{C}) \\ \Delta_C &:= \mathbb{P}(\text{C}|\text{A}, \text{Rep}_\text{X}) - \mathbb{P}(\text{C}|\text{A}). \end{aligned} \tag{6}$$

For the proofs, we refer the reader to the numbered theorems, which are given in full in Appendix B.

Crucially, given the structure of the Bayesian Network, hearing a conditional does *not* change the probability of the antecedent: that is, $\Delta_A = 0$ (see Theorem 1). But on hearing a conditional the probability of the consequent can stay the same, increase, or decrease, in some conditions. We refer readers to footnotes and to Appendix B for a precise specification of the conditions under which change is possible (see, in particular, Theorem 2 and related discussion). Here, we refer readers to the specification of probabilities under (4) and, particularly, the probabilities α β, and γ[32] and note that (i) $\beta \approx \gamma$ and (ii) $\alpha \approx 1$ are jointly sufficient for $\Delta_B \approx 0$.

4.3.1. Summary of the Baseline Model

Across all of the experiments, there was a crucial effect: that the conditional probability increased in response to the assertion of a conditional. The Baseline Model captures this effect: assertion always increases the conditional probability (Theorem 3). As a baseline, the model above can account for some of our data (the first four experiments) under certain conditions. We do not explore whether these conditions actually hold among participants, however, because the model has a serious limitation: The model cannot account for the influence of the priors. We turn, now, to modifications of the baseline model that aim at capturing this.

4.4. The modified Model A

We saw cases in our data where the probability of the antecedent increases: being told “if it rains today, you will get wet” may increase one’s belief in rain, but only where one’s prior is presently low, that is, $\Delta_A \neq 0$. To account for these cases, we modify the baseline model by introducing an additional arc from A to Rep_X. We assign the same probabilities as in Eqs. (3) and (4), but replace Eqs. (5) by the following assignments:

$$\begin{aligned} \mathbb{P}(\text{Rep}_\text{X}|\text{X, Rel, A}) = 1 \quad &, \quad \mathbb{P}(\text{Rep}_\text{X}|\neg\text{X, Rel, A}) = 0 \\ \mathbb{P}(\text{Rep}_\text{X}|\text{X, Rel}, \neg\text{A}) = 0 \quad &, \quad \mathbb{P}(\text{Rep}_\text{X}|\neg\text{X, Rel}, \neg\text{A}) = 0 \\ \mathbb{P}(\text{Rep}_\text{X}|\text{X}, \neg\text{Rel, A}) = \mu \quad &, \quad \mathbb{P}(\text{Rep}_\text{X}|\neg\text{X}, \neg\text{Rel, A}) = \mu \\ \mathbb{P}(\text{Rep}_\text{X}|\text{X}, \neg\text{Rel}, \neg\text{A}) = \mu \quad &, \quad \mathbb{P}(\text{Rep}_\text{X}|\neg\text{X}, \neg\text{Rel}, \neg\text{A}) = \mu \end{aligned} \tag{7}$$

The crucial modification is that we set $\mathbb{P}(\text{Rep}_\text{X}|\text{X, Rel}, \neg\text{A}) = 0$, i.e. if the antecedent is false and if X is true, then a fully reliable information source will not submit the report Rep_X.

We are now in a position to assess changes to the probability of the antecedent, the probability of the consequent, and the conditional probability, defined as in Eqs. (6). Consider the Bayesian Network in Fig. 11 with the probability distribution defined in Eqs. (3), (4) and (7). Firstly, we note that, as before, the conditional probability increases on the assertion of a conditional, that is, $\Delta'_C > 0$ (see Theorem 6). Turning to the probability of the antecedent, Δ'_A, can now stay the same, increase, or decrease. Once again, we refer readers to the Appendix for precise specification of the relevant conditions and, in particular, to Theorem 4. The key point is that, for a wide range of values of the reliability parameter, this model can capture change to the probability of the antecedent.[33] So Modified Model A allows one to capture changes to the probability of the antecedent, at least in principle, where the exact change depends both on the prior and the characteristics of the source.

However, the model fares less well with the probability of the consequent, Δ'_B. On the assertion of a conditional, this probability always increases (see Theorem 5). In the data, change to the probability of the consequent depended on its prior.

4.4.1. Summary of Model A

Model A is an advance on the Baseline Model, then, in allowing change to the probability of the antecedent that is dependent on the prior probability and source characteristics. But in Model A assertion always increases the probability of the consequent, and so the model fails to allow for dependence on the prior that we saw in the data.

4.5. The modified Model B

We now consider whether this dependency on priors can be captured through an alternative modification to the Baseline Model.

[32] Formally, this change is determined by the following conditions. Let $\phi_B := \bar{\alpha}\, a + (\beta - \gamma)\bar{a}$. Then (i) $\Delta_B > 0$ iff $\phi_B > 0$, (ii) $\Delta_B = 0$ iff $\phi_B = 0$, and (iii) $\Delta_B < 0$ iff $\phi_B < 0$. As a consistency check, it is easy to see that $\Delta_B = 0$ if $\mathbb{P}(\text{C}|\text{X}) = a + \beta\,\bar{a} \equiv \mathbb{P}(\text{C}|\text{X}) = \alpha\, a + \gamma\,\bar{a}$ (as it should be). Here and throughout we use the shorthand notation $\bar{x} := 1 - x$ and “iff” stands for the biconditional “if and only if”.

[33] This change is dependent on the following conditions. Let $r_0 := \mu/(\mu + \bar{a})$. Then $\Delta'_A > 0$ iff (i) $r > r_0$, (ii) $\Delta'_A = 0$ iff $r = r_0$, and (iii) $\Delta'_A < 0$ iff $r < r_0$. For small values of a, r_0 is small and hence $\Delta'_A > 0$ for a whole range of values of the reliability parameter r.

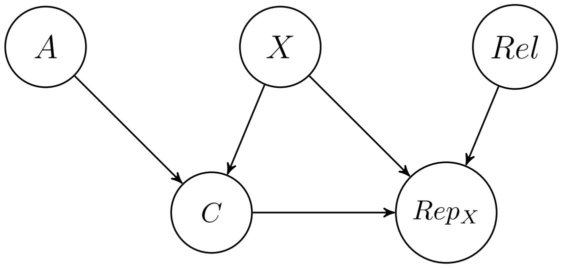

Fig. 12. The Bayesian Network for the Modified Model C.

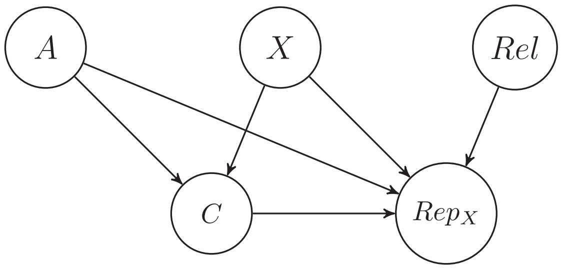

Fig. 13. The Bayesian Network for the Modified Model D.

Instead of Eq. (4) we make the following assignments:

$$\begin{aligned} &\mathbb{P}(\mathrm{C}|\mathrm{A},\mathrm{X}) = \bar{\epsilon} \quad , \quad \mathbb{P}(\mathrm{C}|\mathrm{A},\neg\,\mathrm{X}) = \alpha \\ &\mathbb{P}(\mathrm{C}|\neg\,\mathrm{A},\mathrm{X}) = \beta \quad , \quad \mathbb{P}(\mathrm{C}|\neg\,\mathrm{A},\neg\,\mathrm{X}) = \gamma \end{aligned} \tag{8}$$

with a small parameter ϵ. Contrary to Eqs. (4), we now assume that the consequent B only occurs with high probability (and not with probability 1) if the antecedent holds and if the conditional is true. This might be psychologically more realistic. Considering the Bayesian Network in Fig. 10 and the probability distribution defined in Eqs. (3), (5) and (8), we observe the following.[34] There is no change in the probability of the antecedent, that is, $\Delta_A = 0$ (Theorem 7). The probability of the consequent, in contrast, can stay the same, increase, or decrease (Theorem 8).[35] Lastly, the conditional probability can also stay the same, increase, or decrease (Theorem 9).[36]

4.5.1. *Summary of Model B*

Adding a small error parameter ϵ comes at the cost of ruling out change to the antecedent; hence, at the cost of not capturing an important pattern in our data. In Model B, assertion can also lead to the conditional probability staying the same or even decreasing, whereas our data show a consistent increase in the conditional probability. More data would be needed to test whether the conditional probability might stay the same or even decrease. For present purposes, it suffices to note that, like the preceding models, this model cannot capture all the patterns in our data.

4.6. *The modified Model C*

In this section, we study another extension of the Baseline Model: we introduce an additional arc from C to Rep_X. We assign the same probabilities as in Eqs. (3) and (4), but replace Eqs. (5) by the following assignments:

$$\begin{aligned} &\mathbb{P}(\mathrm{Rep}_\mathrm{X}|\mathrm{X},\mathrm{Rel},\mathrm{C}) = 1 \quad , \quad \mathbb{P}(\mathrm{Rep}_\mathrm{X}|\neg\,\mathrm{X},\mathrm{Rel},\mathrm{C}) = 0 \\ &\mathbb{P}(\mathrm{Rep}_\mathrm{X}|\mathrm{X},\mathrm{Rel},\neg\,\mathrm{C}) = 0 \quad , \quad \mathbb{P}(\mathrm{Rep}_\mathrm{X}|\neg\,\mathrm{X},\mathrm{Rel},\neg\,\mathrm{C}) = 0 \\ &\mathbb{P}(\mathrm{Rep}_\mathrm{X}|\mathrm{X},\neg\,\mathrm{Rel},\mathrm{C}) = \mu \quad , \quad \mathbb{P}(\mathrm{Rep}_\mathrm{X}|\neg\,\mathrm{X},\neg\,\mathrm{Rel},\mathrm{C}) = \mu \\ &\mathbb{P}(\mathrm{Rep}_\mathrm{X}|\mathrm{X},\neg\,\mathrm{Rel},\neg\,\mathrm{C}) = \mu \quad , \quad \mathbb{P}(\mathrm{Rep}_\mathrm{X}|\neg\,\mathrm{X},\neg\,\mathrm{Rel},\neg\,\mathrm{C}) = \mu \end{aligned} \tag{9}$$

The crucial modification is that we set $\mathbb{P}(\mathrm{Rep}_\mathrm{X}|\mathrm{X},\mathrm{Rel},\neg\,\mathrm{C}) = 0$, i.e., if the consequent is false and if X is true, then a fully reliable information source will not submit the report Rep_X. Considering the Bayesian Network in Fig. 12 and the probability distribution

[34] Obviously, for $\epsilon = 0$, we recover Theorems 1–3.

[35] The change depends on the following conditions. Let $\widetilde{\phi}_B := (\overline{\alpha} - \epsilon)a + (\beta - \gamma)\overline{a}$. Then (i) $\Delta_B > 0$ iff $\widetilde{\phi}_B > 0$, (ii) $\Delta_B = 0$ iff $\widetilde{\phi}_B = 0$, and (iii) $\Delta_B < 0$ iff $\widetilde{\phi}_B < 0$.

[36] This change depends on the relationship between the parameters $\overline{\alpha}$ and ϵ. The conditional probability (i) increases ($\Delta_C > 0$) if and only if $\overline{\alpha} > \epsilon$, (ii) stays the same ($\Delta_C = 0$) if and only if $\overline{\alpha} = \epsilon$, and (iii) decreases ($\Delta_C < 0$) if and only if $\overline{\alpha} < \epsilon$.

defined in Eqs. (3), (4) and (9), we observe the following. The probability of the antecedent always increases, that is, $\Delta_A > 0$ (Theorem 10). So, too, does the probability of the consequent, that is, $\Delta_B > 0$ (Theorem 11); and, likewise, the conditional probability, that is, $\Delta_C > 0$ (Theorem 12).

4.6.1. Summary of Model C

In Model C, adding an additional arc from C to Rep_X means that assertion of the conditional leads to an increase in the conditional probability. This is an aspect of the data that was lost in Model B. But assertion also always leads to increases in the probability of the antecedent and the probability of the consequent. Although such increases were seen in our data, there was a dependency on priors that this model cannot capture.

4.7. The modified Model D

Finally, we study a model that introduces an additional arc from A and C to Rep_X. This model can be seen as a combination of Models A and C. We assign the same probabilities as in Eqs. (3) and (4), but replace Eqs. (5) by the following assignments:

$$\begin{aligned}
\mathbb{P}(\mathrm{Rep_X}|\mathrm{X, Rel, A, C}) = 1 \quad &, \quad \mathbb{P}(\mathrm{Rep_X}|\neg \mathrm{X, Rel, A, C}) = 0\\
\mathbb{P}(\mathrm{Rep_X}|\mathrm{X, Rel, A}, \neg \mathrm{C}) = 0 \quad &, \quad \mathbb{P}(\mathrm{Rep_X}|\neg \mathrm{X, Rel, A}, \neg \mathrm{C}) = 0\\
\mathbb{P}(\mathrm{Rep_X}|\mathrm{X, Rel}, \neg \mathrm{A, C}) = 0 \quad &, \quad \mathbb{P}(\mathrm{Rep_X}|\neg \mathrm{X, Rel}, \neg \mathrm{A, C}) = 0\\
\mathbb{P}(\mathrm{Rep_X}|\mathrm{X, Rel}, \neg \mathrm{A}, \neg \mathrm{C}) = 0 \quad &, \quad \mathbb{P}(\mathrm{Rep_X}|\neg \mathrm{X, Rel}, \neg \mathrm{A}, \neg \mathrm{C}) = 0\\
\mathbb{P}(\mathrm{Rep_X}|\mathrm{X}, \neg \mathrm{Rel, A, C}) = \mu \quad &, \quad \mathbb{P}(\mathrm{Rep_X}|\neg \mathrm{X}, \neg \mathrm{Rel, A, C}) = \mu\\
\mathbb{P}(\mathrm{Rep_X}|\mathrm{X}, \neg \mathrm{Rel, A}, \neg \mathrm{C}) = \mu \quad &, \quad \mathbb{P}(\mathrm{Rep_X}|\neg \mathrm{X}, \neg \mathrm{Rel, A}, \neg \mathrm{C}) = \mu\\
\mathbb{P}(\mathrm{Rep_X}|\mathrm{X}, \neg \mathrm{Rel}, \neg \mathrm{A, C}) = \mu \quad &, \quad \mathbb{P}(\mathrm{Rep_X}|\neg \mathrm{X}, \neg \mathrm{Rel}, \neg \mathrm{A, C}) = \mu\\
\mathbb{P}(\mathrm{Rep_X}|\mathrm{X}, \neg \mathrm{Rel}, \neg \mathrm{A}, \neg \mathrm{C}) = \mu \quad &, \quad \mathbb{P}(\mathrm{Rep_X}|\neg \mathrm{X}, \neg \mathrm{Rel}, \neg \mathrm{A}, \neg \mathrm{C}) = \mu
\end{aligned} \tag{10}$$

The crucial modification is that for a reliable source only $\mathbb{P}(\mathrm{Rep_X}|\mathrm{X, Rel, A, C}) = 1$. All other conditional probabilities of Rep_X vanish. Considering the Bayesian Network in Fig. 13 with the probability distribution defined in Eqs. (3), (4) and (10), we observe the following. The probability of the antecedent always increases, that is, $\Delta_A > 0$ (Theorem 13); as does the probability of the consequent, that is, $\Delta_B > 0$ (Theorem 14); and the conditional probability, that is, $\Delta_C > 0$ (Theorem 15).

4.7.1. Summary of Model D

Adding an extra from nodes A and C to Rep_X results in the same pattern as Model C: an increase in all probabilities and no dependency on the priors. Hence this model, too, fails to capture the data.

4.8. Summary of modeling

How well, then, do these models do in accounting for the data? As we have seen, the explicit models of the information-gathering process (the testimonial context) we constructed have more success than the distance-based approach to Bayesianism (which does not take the information-gathering into account). But each model only ever matches some trends within the data. Of the five models we describe, only Model A shows the required dependency on prior probability, but only for the probability of the antecedent.

Even though we have only considered model fit in a loose, qualitative sense, we can reasonably conclude that this method, though an apparently plausible extension of the probabilistic approach to the conditional, fails to capture our simple data set.

5. The challenge to other accounts of conditionals

We began with an assumption that the Suppositional Theory is a good starting point for developing an account of learning from a conditional testimony. While it itself does not provide us with an explanation of how hearers accommodate new conditional information, the massive empirical support, even if qualified, for the probabilistic interpretation of the conditional, motivated our own model. Our data, however, pose a challenge to the Suppositional Theory of the conditional, and its extension, and one could wonder if these negative results do not undermine the Bayesian paradigm entirely. In this section we show that they do not only pose a challenge for this account. Indeed, we will demonstrate in this section that the data pose a broader challenge by considering possible worlds semantics, the material conditional account, and Mental Models Theory. The table below summarizes, once more, the constraints our empirical results put on any accounts of indicative conditionals that aspires to descriptive accuracy.

5.1. The possible worlds account

As outlined earlier, the Suppositional Theory was inspired by the idea of the Ramsey Test, which seems to capture our intuitions about the mental procedure we follow when evaluating indicative conditionals: when deciding whether to accept a conditional, we suppose its antecedent and then decide on the basis of that supposition whether to accept the consequent. The probabilistic Suppositional Theory is, however, just one way of formalizing this idea. An alternative theory of conditionals, which also aims at capturing the intuition expressed by Ramsey, has been developed by Stalnaker (1968, 1975) and, independently, by Lewis (1973).

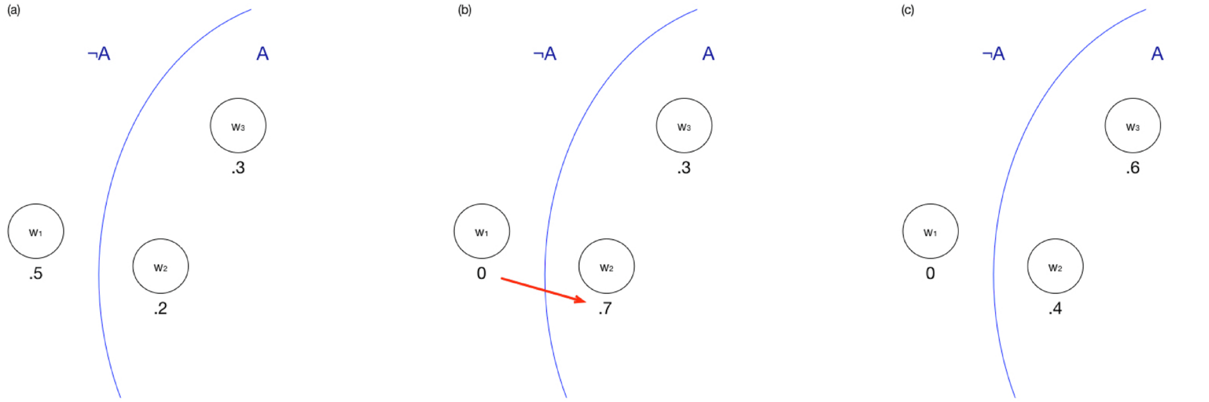

Fig. 14. The Figure shows a simple possible worlds model, consisting of three words and representing two propositions, A and its negation, ¬A. Worlds w_2 and w_3 belong to the proposition A, are A-worlds, while w_1 is a ¬A-world. The world w_2 is the nearest possible A-world accessible from w_1. Figure (a) shows the prior probability assignment. Figures (b) and (c) illustrate the effect of updating on A: (b) shows the new probability assignment after imaging on A, and (c) shows the new probability assignment after conditionalizing on A.

Since Lewis's account was meant as a semantics for counterfactuals, not indicatives, and it was Stalnaker who made an explicit appeal to the Ramsey Test, we will follow Stalnaker's presentation. It should be noted, nevertheless, that the differences between the two accounts of conditionals lie mainly in the details, and thus it is justified—and customary—to treat them as one approach.[37]

In possible worlds semantics, a proposition expressed by a sentence P is a set of possible worlds in which P is true. Possible worlds are to be understood, roughly, as different ways our actual world might be or might have been, that is, different alternatives to our world. They correspond to what we imagine when we suppose that things were different than they actually are, or when we suppose things to be in a particular way when we don't know what is actually true.[38] Take, for instance, the sentence "Berlin is the capital of Germany." We can imagine, for instance, that Berlin is the capital of Germany while San Francisco is the capital of the United States, or while Seattle is the capital of the United States, or while it is Washington, D.C., and so on. We can imagine Berlin being the capital of Germany when there are no other countries on the planet. All these possible, alternative scenarios or descriptions of the world are what semanticists call possible worlds belonging to the proposition expressed by "Berlin is the capital of Germany." But the capital of Germany might have been Munich, or it might have been Düsseldorf, or it might have been Amsterdam, and so on. These are the ways "Berlin is the capital of Germany" might have been false, that is, imagining the capital of Germany being Munich, Düsseldorf, or Amsterdam, corresponds to considering different possible worlds in which the sentence is false. Moreover, we might find some of those ways the world might have been different more or less plausible. An imaginary world in which the capital of Germany is Munich seems more plausible, more similar to the actual world, than a world in which it is Amsterdam. We may also imagine things whose truth value we don't (yet) know, and order them according to their plausibility. For instance, we might imagine a world in which California becomes independent from the United States by the year 2050, even if we find it more plausible that it will not.[39]

On the possible worlds account, conditionals express propositions, which means that they can be true or false. But the truth value of a conditional is not simply a function of the truth values of its parts, i.e., possible worlds semantics is not truth-functional. It is, however, truth-conditional, that is, one can define the conditions under which conditionals are true. Evaluation of a conditional, "If A, then C," on Stalnaker's account, begins, as suggested by Ramsey, with a *supposition* of the antecedent, that is, with imagining the nearest possible world (or worlds, on Lewis' version of the theory) in which the antecedent, A, is true (the nearest A-worlds, for short). This means imagining the world as it actually is except for what needs to be different for A to be true. For instance, to evaluate "If California becomes an independent country, then the United States will be a federation of 49 states," we imagine that California declared independence, and perhaps that it is recognized as independent by other states and so on, but everything else that we know about our actual world remains fixed. The worlds in which Alabama is also an independent country, or the worlds in which the states are not united at all, are *not* the nearest possible worlds in which "California becomes an independent country" is true. Once we fix the nearest antecedent world, we evaluate the consequent of the conditional in that world. More formally, Stalnaker's conditional, denoted by $A > C$, is true in the actual world if and only if C is true in the nearest possible A-world (or, on Lewis account, in all nearest A-worlds). A proposition expressed by $A > C$ is then the collection of all worlds in which $A > C$ is true. There are two types of relevant worlds, namely, all those worlds in which both A and C are true and all those ¬A-worlds such that C is true in their nearest A-worlds. Since how we order the possible worlds, that is, which worlds we consider to be close to each other, depends on our epistemic states (e.g. our background knowledge), the same conditional may express different propositions for different speakers in different

[37] Yet another related account has been proposed by Kratzer (1979, 1986).

[38] The notion of a possible world, although primarily a formal construct, is not devoid of psychological reality. Rips and Marcus (1977), for instance, discuss suppositions as cognitive analogues of the possible worlds. The notion of a possible world also proved to be useful in the context of developmental research on counterfactual reasoning (see, e.g., Rafetseder, Cristi-Vargas, & Perner, 2010; Leahy, Rafetseder, & Perner, 2014).

[39] There is psychological evidence that people often have clear intuitions about which alternatives to the reality are more plausible than others. See, e.g., Kahneman and Miller (1986), Byrne (2005), Perner and Rafetseder (2011), Markman, Gavanski, Sherman, and McMullen (1993).

contexts (cf. Stalnaker, 1984, p. 109).

What would happen, then, if people updated their beliefs on a Stalnaker conditional? Since conditionals on Stalnaker's account express propositions, the first step in modeling learning from a conditional seems straightforward: we fix the posterior probability of a conditional to a new, higher value. When the source of the testimony is reliable, we assign a higher probability value to the conditional than when the source is not fully reliable. What, then, is the conditional probability of the consequent given the antecedent? Stalnaker himself proposed that the probability of a conditional should equal the corresponding conditional probability.[40] For convenience we repeat, here, Stalnaker's formulation from above:

Stalnaker's Hypothesis ("The Equation"): For all probability functions $\mathbb{P}$ such that $\mathbb{P}(A) > 0$, $\mathbb{P}(\text{If A, C}) = \mathbb{P}(C|A)$.

However, Stalnaker's semantics does not validate "the Equation." In fact, as proven by Lewis (1976), $\mathbb{P}(A > C)$ is the probability of C, not after conditioning on A, but after *imaging* on A, denoted by $\mathbb{P}(C)[A]$.[41] Intuitively, updating by imaging on A amounts to taking the probability of each ¬A-world, and moving it to its closest A-world. As Fig. 14 illustrates, none of the ¬A-worlds keeps its original probability, but all A-worlds do, and, additionally, some A-worlds become more probable when an additional share of the probability is shifted from some $\neg A$-worlds. The outcome of Imaging depends then on how we order the possible worlds, which might vary depending on the context of an utterance or on our epistemic states. For instance, if the model shown on Fig. 14 were such that w_3 was closer, i.e., more similar to w_1 than w_2, in (b) we would have transfered the probability assigned to w_1 to w_3, and thus the posterior probability of w_3 would have been .8, while the probability assigned to w_2 would have stayed the same, that is, it would have remained .2. At the same time, ordering of the possible worlds has no effect on the Bayesian updating, and hence the posterior probability distribution would have been the same (c).[42]

While the probability of a Stalnaker's conditional itself is given by imaging, one can use the conditional to update other beliefs via conditionialization (as the conditional expresses a proposition). Consequently, an agent learning a conditional can, in principle, update either by classical conditionalization or by imaging.[43] In both cases, the conditional probability of C given A increases; however, the effect on the posterior probability of A and C is less straightforward. Depending on the possible worlds model we construct, which in turn might depend on the context in which the conditional is asserted, learning a conditional can result in an increase or a decrease of the probability of a proposition, or it can have no effect on it at all. Although this might seem to be an advantage, as observed by Douven (2012, p. 246–247), such flexibility does not mean that the posterior probability assignments resulting from either imaging or conditionalization on a conditional will correspond to our intuitions about any particular case.[44] To account for the outcomes of the experiments reported above, Stalnaker's semantics would need to be accompanied by an auxiliary (possibly pragmatic) account that would not only explain how people construct their similarity orderings and how they choose the update rule, but also, and most importantly, why they change their beliefs when the semantics does not predict that they would: given that at the heart of the Stalnaker conditional is supposition ("what would the world be like for C if A were true"), it is unclear why one's beliefs about the conditional itself, or A, or C should change at all. Without a systematic account to that effect, however, relegating the explanation of the unpredicted changes to broadly construed pragmatics may strike one as an *ad hoc* solution.

5.2. The material conditional

We have so far considered theories of the conditional which aim at capturing the mental process involved in the evaluation of the conditional known as the Ramsey Test. We turn, now, to two theories that start from a different angle and take classical propositional logic as their main inspiration. Take, first, the material interpretation of the conditional. Although the material conditional interpretation of the conditional has lost popularity in philosophy (though see Rieger (2013) for a recent defense of the material interpretation), it has arguably had a lasting influence on the psychology of reasoning, as we will see below. The material conditional, $A \supset C$, is true if and only if either A is false or C is true (in other words $\neg A \vee C$). Since the material conditional is truth functional, and hence its truth value does not depend on what its antecedent and consequent are actually about, it is prone to well-known counterintuitive consequences, such as the paradoxes of material implication: the material conditional renders true both conditionals whose antecedent and consequent are true but entirely unconnected such as "If kangaroos are mammals, Paris is the capital of France" and, even worse, all conditionals whose antecedents are false such as "If Paris is the capital of Belgium, then Paris is the capital of France." This is because a true consequent or a false antecedent suffices to establish the truth value of a conditional; the content of the remaining clauses has no bearing on the evaluation of the whole sentence. To explain the divergence from our intuitions, the material

[40] Note, however, that, unlike the Suppositional Theory, Stalnaker's theory faces the challenge of the Triviality Results (Lewis, 1976). We will not discuss this issue here.

[41] For experimental investigations of learning a categorical information via imaging, see, e.g., Baratgin and Politzer (2010) and Zhao and Osherson (2014). See also Lucas and Kemp (2015) for a discussion of imaginig on counterfactual conditionals in causal Bayes nets.

[42] One of the consequences of the fact that the probability of a Stalnaker's conditional is obtained by imaging is that its probability and the conditional probability of its consequent given the antecedent can come apart: the conditional probability can be high when the conditional is not very probable, and vice versa.

[43] Or by Jeffrey's version of either of these update rules. For a recent model of learning a Stalnaker's conditional by Jeffrey Imaging, see Günther (2017).

[44] Douven (2012, p. 246–247) presents three models that differ only in how the possible worlds are ordered to show that, first, this is enough for the Stalnaker's conditional to express a different proposition in each of the models, and that both imaging and classical conditionalization on those propositions result in different posterior probability distributions.

conditional account is usually supplemented by a pragmatic account allowing separation of conditionals' truth conditions and their assertability conditions (Grice, 1989; Jackson, 1979).

If a natural language conditional is interpreted as $A \supset C$, learning that "if A, then C" can be modeled as updating by means of Classical Conditionalization on a proposition expressed by a disjunction, $\neg A \vee C$. For any proposition B in the algebra over which the probability measure $\mathbb{P}$ is defined, we can now calculate its posterior probability as $\mathbb{P}(B| \neg A \vee C)$. What happens to the probability of the antecedent and the probability of the consequent on such an account? Popper and Miller (1983) observed the following fact (see Douven, 2012):

Popper and Miller (1983): For all A, C, such that $\mathbb{P}(A) < 1$ or $\mathbb{P}(C) < 1$ (or both), $\mathbb{P}(A|A \supset C) \leqslant \mathbb{P}(A)$. For all A and C such that $0 < \mathbb{P}(A) < 1$ and $\mathbb{P}(C|A) < 1$, $\mathbb{P}(A|A \supset C) < \mathbb{P}(A)$.

This means that, as long as our prior $\mathbb{P}(A)$ is greater than 0 and less than 1, learning that $A \supset C$ should make us decrease our degrees of belief in A. At the same time, as long as our prior $\mathbb{P}(C) < 1$ and greater than 0, we should become more confident that C is the case. The reader will recall that these are exactly the predictions made by the distance-based approach (which is equivalent to conditionalizing on the material conditional if the conditional is strict, i.e. if one takes $\mathbb{P}'(C|A) = 1$ as a constraint on $\mathbb{P}'$).

As we saw for the distance-based account, our data do not support this pattern of belief change. The data suggest a dependence on prior probabilities, and this dependence has no place in the material account. By contrast, for $\mathbb{P}(A) > 0$ and $\mathbb{P}(C) < 1$, our $\mathbb{P}'(C|A)$ should be higher than our prior $\mathbb{P}(C|A)$, which is exactly what we expected. Note, however, that the probability of a conditional, $\mathbb{P}(A \supset C)$, does not equal the corresponding conditional probability, $\mathbb{P}(C|A)$, since it is a theorem of the probability calculus that:

$$\mathbb{P}(A \supset C) = \mathbb{P}(\neg A) + \mathbb{P}(A) \cdot \mathbb{P}(C|A). \tag{11}$$

Furthermore, it is easy to see from Eq. (11) that the less likely the antecedent, the more likely the conditional, represented by the material conditional, provided that the conditional probability $\mathbb{P}(C|A)$ does not change (cf. Douven, 2016a, p. 57). Eq. (11) also shows that $\mathbb{P}(A \supset C) = 1$ if and only if $\mathbb{P}(C|A) = 1$ (provided that $\mathbb{P}(A)$ is non-extreme). This fact and the result mentioned above that Classical Conditionalization follows from the minimization of an *f*-divergence explains why conditionalizing on the material conditional and the minimization of an *f*-divergence are equivalent if the learned conditional is strict, i.e., if one takes $\mathbb{P}'(C|A) = 1$ as a constraint on $\mathbb{P}'$. At the same time, it is easy to see that $\mathbb{P}(C|A) = 1 - \epsilon$ does not imply that $\mathbb{P}(A \supset C)$ is also close to 1. That is, if the conditional is non-extreme, then we would expect, in general, the material conditional plus Jeffrey Conditionalization and the conditional probability approach to make different predictions (see Eq. (11)). This parallels the counter-intuitive effects of introducing a parameter ϵ to allow $\mathbb{P}(C|A) = 1 - \epsilon$ in Modified Model B above. For further discussion, see Eva et al. (2019).

However, the material conditional can, at least, deal with our reliability manipulation. Since a material conditional expresses a proposition which can be represented by the set of all those worlds in which the disjunction $\neg A \vee C$ is true, one can model the effect of the reliability manipulations in a straightforward way, namely, by assigning a higher probability value (for instance 1) to the proposition $\neg A \vee C$ when the conditional comes from a reliable source, and a lower probability value when the source is not reliable. The account provides no guidance as to what these values should actually be, but it is, at least, qualitatively compatible. This strategy should work for any account on which an indicative conditional expresses a proposition.

5.3. Mental models theory

In the psychology of reasoning, although the material conditional account has been long since rejected as an interpretation of natural language conditionals, it has nevertheless influenced a vast amount of reasoning research in its early days. These early reasoning experiments contributed to the development of one of the leading theories in the psychology of reasoning, the Mental Models Theory (MMT) (Johnson-Laird & Byrne, 2002; Johnson-Laird et al., 2015; Khemlani, Byrne, & Johnson-Laird, 2018). On this theory, the *core meaning* of an indicative conditional refers to the conjunction of the following possibilities:

A	C
¬A	¬C
¬A	C

An interpretation of a conditional amounts to imagining (not imaging!) the possibilities (referred to as "mental models") that are compatible with the conditional: the possibility that both A and C are the case, the possibility that neither A nor C are the case, and the possibility that A is not the case but C is the case. When interpreting a conditional, however, people do not always imagine the fully explicit model. They may stop, for instance, already after representing the possibility that A and C:

A	C
…	

The ellipsis indicates that the model did not make all the relevant possibilities explicit. Such a mental model is what underlies intuitive reasoning, although the fully fleshed out explicit models can be evoked if the tasks are simple enough (e.g. Khemlani et al., 2018, p. 9).

It is important to emphasise that, although there seems to be a correspondence between the set of possibilities that comprise the core meaning of the conditional, on the one hand, and the set of cells in a truth table for the material conditional in which the conditional is true, the logic of a conditional on MMT is not that of the material conditional. On the most recent version of the MMT (Johnson-Laird et al., 2015; Hinterecker, Knauff, & Johnson-Laird, 2016; Khemlani, Lotstein, & Johnson-Laird, 2015; Khemlani et al., 2018), it is not only the case that the conditional is compatible with the three mental models shown above, but it is true only if all the three situations represented by the fully explicit model are possible. That is, "If A, then C" is true, on the MMT, if and only if it is possible that A and C, *and* it is possible that ¬A and ¬C, *and* it is possible that ¬A and C. In other words, the conditional is equivalent to the following formula (where $\Diamond$ is read as "possibly"):

$$\Diamond(A \wedge C) \wedge \Diamond(\neg A \wedge C) \wedge \Diamond(\neg A \wedge \neg C) \tag{12}$$

Additionally, given that the conjunction of possibilities is supposed to be interpreted as exhaustive, one can infer from it that A and ¬C is impossible: $\neg\Diamond(A \wedge \neg C)$. Consequently, the revised MMT seems to predict that what an agent learns when they hear a conditional is precisely (12), and, implicitly, the impossibility of A and ¬C.

How should an agent adjust their degrees of belief upon hearing a conditional, if that conditional is indeed represented as the MMT predicts? To account for the growing body of data on probabilistic reasoning, the initially qualitative Mental Models Theory has been extended by the theory of "naive probability" (Khemlani et al., 2015; Johnson-Laird et al., 2015; Hinterecker et al., 2016), which is a dual-process account of probabilistic reasoning. On this account, the slow and deliberative System 2 tends to comply with Bayesian probability theory, while the fast, intuitive System 1 gives estimates based on certain simple heuristics. More specifically, Khemlani et al. (2015) provide an algorithm for estimating intuitive probabilities of unique events and their compounds: conjunctions, disjunctions, and conditional probability. The latter, $\mathbb{P}(C|A)$ is computed as the ratio of $\mathbb{P}(A \wedge C)$ to $\mathbb{P}(A)$ by the deliberative System 2.[45] If only System 1 is employed, however, the conditional probability is said to be estimated by adjusting the probability of C, $\mathbb{P}(C)$, depending on whether the antecedent, A, increases or decreases it (Johnson-Laird et al., 2015, p. 211). In other words, the intuitive System 1, on this account, treats A as evidence for the estimates of the probability of C, and that probability becomes the conditional probability $\mathbb{P}(C|A)$ (Khemlani et al., 2015, p. 1233).

Interestingly, Khemlani et al. (2015) do not say anything about estimating the probabilities of natural language conditionals and how these relate to the corresponding conditional probabilities, given that conditionals are represented as conjunctions of (epistemic) possibilities. It is difficult to see how the probability of (12) could even approximate the conditional probability of C given A, and, how it could have various degrees of probability less than 1 unless there is still a possibility, even if highly unlikely, that A and ¬C is the case.[46] Consequently, it is unclear how the MMT could account for different conditional probability assignments depending on the reliability of the speaker.[47]

There is however one finding with respect to which the new MMT might be used to make a (qualitatively) correct prediction, namely, that the probabilities of antecedents and consequents change in response to learning the conditional in the situation in which their prior probabilities are extremely low. For instance, if we believe that it is highly unlikely that Lisa, a Liberal Arts College, is majoring in astrophysics, accepting the testimony "If Lisa, a student, is majoring in astrophysics, then she's working late in the library" forces us to represent the situation in which Lisa is majoring in astrophysics and in which she's working late in the library as possible. Once such a possibility is brought to our attention, we might indeed find it more probable that Lisa is indeed majoring in astrophysics.

Finally, an important addition to the MMT are mechanisms of semantic and pragmatic modulations that are said to affect what we find possible or impossible, and how the mental models are constructed (Johnson-Laird & Byrne, 2002; Quelhas, Johnson-Laird, & Juhos, 2010; Khemlani et al., 2018). Perhaps, then, MMT proponents might invoke these modulations to try and reconcile their theory with our findings on how people change their beliefs upon learning a conditional. However, unless developed into a full-fledged, systematic account, such a move risks being an *ad hoc* solution.

6. Conclusions

We presented a number of simple, fundamental intuitions about people's responses to the testimonial assertion of a conditional: that hearing a source assert a conditional will lead to in an increase in the conditional probability of the consequent, given the antecedent (at least for conditional relationships that we do not yet have full knowledge of). At the same time, that increase is modulated by the reliability of the source. And, finally, assertion of a conditional can alter our beliefs about the probability of the antecedent and the consequent: raising that probability when we took it to be low, and potentially lowering it by casting doubt ("if…") when it is high. These intuitions were largely confirmed in behavioral experiments. Despite the fact that the uncertainty involved seems to lend itself naturally to a probabilistic treatment, we showed that once one takes the details of such a treatment seriously, the basic data stubbornly resist Bayesian conditionalization, Jeffrey conditionalization, imaging and a distance-based approach to Bayesian learning. In order to capture the basic feature of testimony, namely that perceived reliability of the source matters, one needs to model that aspect of testimony explicitly. In other words, to capture fully the effect of a conditional on our beliefs, one needs to include aspects of the data gathering process by which we have come to 'know' that assertion. However, we have seen the failure to capture our data of even a common model of testimony, based on a simple Bayesian belief network, that has seen

[45] Khemlani et al. (2015) points out, however, that individuals "have no deliberative method for computing conditional probabilities about unique events" (p. 1233).

[46] We owe this observation to a Reviewer for this journal.

[47] On various problems following from the new revised version of the MMT, see Baratgin et al. (2015) or Oaksford, Over, and Cruz (2019).

considerable theoretical and empirical empirical success (Bovens, Fitelson, Hartmann, & Snyder, 2002; Bovens & Hartmann, 2002, 2003; Collins, Hahn, von Gerber, & Olsson, 2018; Hahn, Harris, & Corner, 2009; Hahn, Oaksford, & Harris, 2012; Harris, Hahn, Madsen, & Hsu, 2016; Jarvstad & Hahn, 2011; Landes, Osimani, & Poellinger, 2018). This model fails to adequately capture our data even with multiple modifications. Hence, a very simple phenomenon raises unexpected technical challenges and complexities on closer inspection.

In short, this paper has asked a superficially simple question: how do we change our beliefs when we hear (read) the assertion of a conditional by a partially reliable source? A set of simple experiments delivered apparently straightforward and intuitive results. But these data are enough to challenge major theories of the conditional. We have demonstrated the challenges for the Suppositional Theory (fleshed out both with a distance-based approach and with Bayesian belief networks); the Stalnaker conditional; the material conditional; and Mental Models Theory.

We are left with a paradox and a dilemma. One of the most basic things we could ask about conditionals, namely how we modify our beliefs upon hearing one, seems stubbornly outside the remit of present theories of the conditional. Despite the fact that people find the task of reporting their (new) beliefs on having heard a conditional straightforward and unremarkable, no present theoretical account of the conditional accommodates the very intuitive data patterns that ensue. Moreover, trying to model these data, on any present account found in the literature, raises thorny conceptual questions about what the conditional actually is.

In recent decades, research on conditional reasoning in psychology has made progress in part precisely because it has tried to sidestep questions about the semantics of conditionals. Researchers such as Oaksford and Chater (1994) and following have tried to commit only to the idea that uttering a conditional indicates that the conditional probability is high. In a similar spirit, we have sought to remain agnostic about the meaning of the conditional in the probabilistic models presented in Section "Modeling the Data," assuming only that the variable *X* represents the indicative conditional 'If A, then B.' Nevertheless, semantic tensions arise about the meaning of this variable, and the fact that it would seem to, in some sense, refer to a link within the Bayesian Network (so implying a meta-statement within an object level representation). Our account, outlined in Appendix C, was the best attempt to render this approach semantically well-formed. The reader may or may not find this attempt fully convincing. The difficulties raised, however, apply not just to the specific models presented, but to any probabilistic account that would seek to use the machinery of belief revision within a Bayesian Network as its basis. A broader probabilistic framework thus seems required, but those considered (imaging and the distance-based approach without an explicit consideration of the reliability of the source) sit uneasily with the data. Of course, new probabilistic methods for the transition between models may be found, which capture better the change to beliefs prompted by hearing a conditional. However, we know of no obvious further route to pursue.

We have collected only simple data on how beliefs change in response to a conditional, and yet these data pose a considerable challenge. This fact seems to have profound implications for the study of conditional reasoning, which has long been the focal point of research on conditionals. In particular, conditional reasoning seems rather more poorly understood than previously recognized. If an account of conditional reasoning starts only at the point where a conditional is magically in place, it is arguably missing a key part of what we do with conditionals in everyday cognition.

As mentioned earlier, it is arguably the historical focus of classical logic that has allowed this fundamental gap to be missed. Classical logic concerns itself with what follows from given premises, not the truth or falsity of those premises as such. Consequently, the question of how we come to know those premises is outside of the focus of theoretical attention. A key contribution of the present paper is to highlight how a complete account of conditional reasoning must involve not just an account of where those premises come from, but also an account that can fit with how we reason from whatever it is we have acquired.

This also poses a challenge for potential routes beyond the framework of either probability or logic. One might, for example, be tempted to move beyond normative frameworks, such as probability theory or logic, and try to seek a purely descriptive account of how we learn conditionals (Elqayam & Evans, 2011). However, a primary interest in conditionals is in the inferences they license, which is a normative question. The challenge is thus a mirror image to that just highlighted for classical logic. A theory of what is learned must be able to interface meaningfully with a theory of subsequent inference. In this case, a purely descriptive cognitive model must provide the right building blocks for a follow-on theory of inference and implication.

Finally, this "interface problem" is also relevant to a final route readers may be tempted to pursue. On this route, which we referred to above, a critic might compartmentalize some of our findings: changes to the probability of the antecedent or the probability of the consequent. Perhaps these changes are merely pragmatics, and should be explained by general principles of conversation; when we build and evaluate our theories of conditionals and conditional reasoning, we can safely ignore them.

We believe this reasoning underestimates the challenge posed by our results. What we have identified is a set of simple, straightforward intuitions about how people change their beliefs on hearing a conditional. These intuitions are supported by experimental data. These straightforward patterns of belief change are part of human dealings with conditionals, and need capturing in a framework that fits with other aspects of belief change – change that arises through subsequent reasoning.

If we are to construct a successful theory of how people respond to the assertion of conditionals, our theory must deliver the right "objects" for subsequent reasoning. By simply pointing to pragmatics, we leave the challenge unaddressed. Our pragmatics needs to be expressed in terms compatible with the rest of our theory of reasoning. Take, for instance, the probabilistic approach. Many pragmatic frameworks are defined over formal objects, such as discourse representation theory (see, for example, Asher & Lascarides, 2003), that do not presently interface with the kind of probabilistic reasoning from conditionals seen in the literature. For the probabilistic approach, a (future) probabilistic pragmatics would, of course, fare better, perhaps of the kind that has become the focus of recent research (Goodman & Lassiter, 2015; Franke & Jäger, 2016). However, this probabilistic pragmatics would need to solve the fundamental problem demonstrated in our modeling sections: the difficulty lies in getting the mechanics of probability theory to work for the straightforward patterns of belief change we identify. Ultimately the same challenge applies to other theories of conditionals and conditional reasoning and their respective mechanics.

In summary, conditionals pervade our daily lives, but a basic understanding of how we use them remains elusive. Decades of multi-disciplinary research efforts have focused on their semantics, the inferences they license, and on descriptive psychological accounts. But these efforts have overlooked the question of what happens when we hear a conditional in the first place, and how our beliefs change in response. This basic question, and the trivial seeming behavioral responses our research reveals, pose a challenge to every theory of the conditional in the literature.

Acknowledgments

For helpful discussion, we thank Igor Douven, Mike Oaksford, Reuben Stern, and the audience of the Learning Conditionals workshop at the Center for Advanced Study, LMU Munich, February 2017. We also thank three anonymous reviewers. Peter Collins was funded by ESRC studentship ES/J500021/1. All authors were funded by awards of the Alexander von Humboldt Foundation to Stephan Hartmann (Alexander von Humboldt Professorship) and Ulrike Hahn (Anneliese Maier Research Award) and through the Senior Researcher in Residence program of LMU's Center for Advanced Studies (awarded to Stephan Hartmann). Gregory Wheeler was also funded by the CoLAForm Project from the ANR-DFG (RO 4548/8-1). Karolina Krzyżanowska was also supported by "The Logic of Conceivability" Project funded by the European Research Council (ERC CoG), Grant No. 681404. We are also grateful for financial support from the Munich Center for Mathematical Philosophy. Peter Collins is now at the Department of Psychology, Social Work, and Counselling, at the University of Greenwich.

Appendix A. Full methods and analysis

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cogpsych.2020.101329.

Appendix B. Theorems and proofs

In this appendix, we first state the theorems mentioned in *Modeling the Data* and then present the proofs.

Theorem 1. *Consider the Bayesian Network in Fig. 10 with the probability distribution defined in Eqs. (3)–(5). Then* $\Delta_A = 0$.

Proof. We note that

$$\mathbb{P}\left(\mathrm{A}|\mathrm{Rep_X}\right) = \frac{\mathbb{P}(\mathrm{A}, \mathrm{Rep_X})}{\mathbb{P}(\mathrm{Rep_X})} \tag{13}$$

and calculate

$$\begin{aligned}\mathbb{P}(\mathrm{Rep_X}) &= \sum_{A,C,Rel,X} \mathbb{P}(A)\mathbb{P}(X)\mathbb{P}(Rel)\mathbb{P}\left(C|A, X\right)\mathbb{P}\left(\mathrm{Rep_X}|X, Rel\right)\\ &= \sum_{Rel,X} \mathbb{P}\left(\mathrm{Rep_X}|X, Rel\right)\mathbb{P}(Rel)\mathbb{P}(X)\\ &= (r + \mu\,\bar{r})x + \mu\,\bar{r}\,\bar{x} = r\,x + \mu\,\bar{r}\end{aligned} \tag{14}$$

Similarly, we obtain:

$$\begin{aligned}\mathbb{P}(\mathrm{A}, \mathrm{Rep_X}) &= \sum_{C,Rel,X} \mathbb{P}(\mathrm{A})\mathbb{P}(X)\mathbb{P}(Rel)\mathbb{P}\left(C|\mathrm{A}, X\right)\mathbb{P}\left(\mathrm{Rep_X}|X, Rel\right)\\ &= a \sum_{Rel,X} \mathbb{P}\left(\mathrm{Rep_X}|X, Rel\right)\mathbb{P}(Rel)\mathbb{P}(X)\\ &= a\,[(r + \mu\,\bar{r})x + \mu\,\bar{r}\,\bar{x}] = a\,(r\,x + \mu\,\bar{r})\end{aligned} \tag{15}$$

Hence, $\mathbb{P}(\mathrm{A}|\mathrm{Rep_X}) = a$ and $\Delta_A = 0$. □

Theorem 2. *Consider the Bayesian Network in Fig. 10 with the probability distribution defined in Eqs. (3)–(5). Furthermore, let* $\phi_B := \bar{\alpha}\,a + (\beta - \gamma)\bar{a}$. *Then (i)* $\Delta_B > 0$ *iff* $\phi_B > 0$, *(ii)* $\Delta_B = 0$ *iff* $\phi_B = 0$, *and (iii)* $\Delta_B < 0$ *iff* $\phi_B < 0$.

Proof. We calculate

$$\begin{aligned}\mathbb{P}(\mathrm{C}, \mathrm{Rep_X}) &= \sum_{A,Rel,X} \mathbb{P}(A)\mathbb{P}(X)\mathbb{P}(Rel)\mathbb{P}\left(\mathrm{C}|A, X\right)\mathbb{P}\left(\mathrm{Rep_X}|X, Rel\right)\\ &= (a + \beta\,\bar{a})(r + \mu\,\bar{r})x + (\alpha\,a + \gamma\,\bar{a})\mu\,\bar{r}\,\bar{x}\end{aligned} \tag{16}$$

and

$$\begin{aligned}\mathbb{P}(C) &= \sum_{A,Rel,Rep_X,X} \mathbb{P}(A)\mathbb{P}(X)\mathbb{P}(Rel)\mathbb{P}\left(C|A, X\right)\mathbb{P}\left(Rep_X|X, Rel\right)\\ &= \sum_{A,X} \mathbb{P}(A)\mathbb{P}(X)\mathbb{P}\left(C|A, X\right)\\ &= a\,x + \alpha\,a\,\overline{x} + \beta\,\overline{a}\,x + \gamma\,\overline{a}\,\overline{x}\end{aligned} \tag{17}$$

With Eq. (14) we can now calculate $\mathbb{P}(\mathrm{C}|\mathrm{Rep}_\mathrm{X})$ and obtain after some algebra:

$$\Delta_B = \frac{r\,x\,\overline{x}}{r\,x + \mu\,\overline{r}}\cdot\phi_B, \tag{18}$$

with

$$\phi_B := \overline{\alpha}\,a + (\beta - \gamma)\,\overline{a}. \qquad \square \tag{19}$$

As a consistency check, it is easy to see that $\Delta_B = 0$ if $P(\mathrm{B}|\mathrm{X}) = a + \beta\,\overline{a} \equiv P(\mathrm{B}|\mathrm{X}) = \alpha\,a + \gamma\,\overline{a}$ (as it should be).

As a corollary of Theorem 2 we conclude that $\Delta_B \approx 0$ if (i) $\beta \approx \gamma$ and (ii) $\alpha \approx 1$. It is also easy to see that $\Delta_B > 0$ if $\beta > \gamma$.

Theorem 3. *Consider the Bayesian Network in Fig. 10 with the probability distribution defined in Eqs. (3)–(5). Then* $\Delta_C > 0$.

Proof. Similarly as before, we obtain

$$\mathbb{P}(\mathrm{C}|\mathrm{A}) = x + \alpha\,\overline{x} \tag{20}$$

and

$$\mathbb{P}(\mathrm{A}, \mathrm{C}, \mathrm{Rep}_\mathrm{X}) = a\,[r\,x + \mu\,\overline{r}\,(x + \alpha\,\overline{x})]. \tag{21}$$

Using Eq. (15), we finally obtain

$$\Delta_C = \frac{\overline{\alpha}\,r\,x\,\overline{x}}{r\,x + \mu\,\overline{r}}, \tag{22}$$

which is greater than zero. $\square$

Theorem 4. *Consider the Bayesian Network in Fig. 11 with the probability distribution defined in Eqs. (3), (4) and (7). Furthermore, let* $r_0 := \mu/(\mu + \overline{a})$. *Then* $\Delta_A > 0$ *iff* $r > r_0$, *(ii)* $\Delta_A = 0$ *iff* $r = r_0$, *and (iii)* $\Delta_A < 0$ *iff* $r < r_0$.

Proof. We calculate the new expressions for $\mathbb{P}(\mathrm{Rep}_\mathrm{X})$ and $\mathbb{P}(\mathrm{A}, \mathrm{Rep}_\mathrm{X})$ and obtain:

$$\mathbb{P}(\mathrm{Rep}_\mathrm{X}) = a\,r\,x + \mu\,\overline{r} \tag{23}$$

$$\mathbb{P}(\mathrm{A}, \mathrm{Rep}_\mathrm{X}) = a(r\,x + \mu\,\overline{r}\,\overline{x}) \tag{24}$$

Hence,

$$\begin{aligned}\Delta'_A &= \frac{a\,x\,(\overline{a}\,r - \mu\,\overline{r})}{a\,r\,x + \mu\,\overline{r}}\\ &= \frac{a\,x\,(\mu + \overline{a})}{a\,r\,x + \mu\,\overline{r}}\cdot\left(r - r_0\right)\end{aligned} \tag{25}$$

with

$$r_0 := \frac{\mu}{\mu + \overline{a}}, \tag{26}$$

from which the theorem follows. $\square$

Theorem 5. *Consider the Bayesian Network in Fig. 11 with the probability distribution defined in Eqs. (3), (4) and (7). Then* $\Delta_B > 0$.

Proof. We calculate

$$\mathbb{P}(\mathrm{C}, \mathrm{Rep}_\mathrm{X}) = a\,r\,x + \mu\,\overline{r}\cdot\mathbb{P}(\mathrm{C}) \tag{27}$$

and use Eq. (23) to obtain

$$\Delta'_B = \frac{a\,r\,x}{a\,r\,x + \mu\overline{r}}\cdot\mathbb{P}\left(\neg\,\mathrm{C}\right), \tag{28}$$

which is greater than zero.

Note that if $\alpha \approx 1$ and $\beta \approx \gamma$, then $\mathbb{P}(\mathrm{C}) \approx a + \beta\bar{a}$, which is approximately 1 if $\beta \approx 1$. (See the discussion below Theorem 2.) Then $\Delta_B' \approx 0$ as we see in the data. Hence, in the modified model, the additional condition $\beta \approx 1$ has to be satisfied to yield $\Delta_B' \approx 0$ unless $a \approx 1$. □

Theorem 6. *Consider the Bayesian Network in Fig. 11 with the probability distribution defined in Eqs. (3), (4) and (7). Then* $\Delta_C > 0$.

Proof. We calculate

$$\begin{aligned} \mathbb{P}(\mathrm{C}|\mathrm{A}) &= x + \alpha\bar{x} \\ \mathbb{P}(\mathrm{A}, \mathrm{C}, \mathrm{Rep_X}) &= a[rx + \mu\bar{r}(x + \alpha\bar{x})] \end{aligned} \tag{29}$$

and use Eq. (24) to obtain

$$\Delta_C' = \frac{x(\bar{\alpha}r\bar{x} + \mu\bar{r}x + \alpha\mu\bar{r}\bar{x})}{rx + \mu\bar{r}\bar{x}}, \tag{30}$$

which is greater than zero. □

Theorem 7. *Consider the Bayesian Network in Fig. 10 with the probability distribution defined in Eqs. (3), (5) and (8). Then* $\Delta_A = 0$.

Proof. It is easy to see that $\mathbb{P}(\mathrm{A}, \mathrm{Rep_X})$ and $\mathbb{P}(\mathrm{Rep_X})$ do not change compared to the situation in Theorem 1. Hence, $\Delta_A = 0$. □

Theorem 8. *Consider the Bayesian Network in Fig. 10 with the probability distribution defined in Eqs. (3), (5) and (8). Furthermore, let* $\widetilde{\phi}_B := (\bar{\alpha} - \epsilon)a + (\beta - \gamma)\bar{a}$. *Then (i)* $\Delta_B > 0$ *iff* $\widetilde{\phi}_B > 0$, *(ii)* $\Delta_B = 0$ *iff* $\widetilde{\phi}_B = 0$, *and (iii)* $\Delta_B < 0$ *iff* $\widetilde{\phi}_B < 0$.

Proof. We calculate

$$\mathbb{P}(\mathrm{C}, \mathrm{Rep_X}) = (\bar{\epsilon}\, a + \beta\bar{a})(r + \mu\bar{r})x + (\alpha a + \gamma\bar{a})\mu\bar{r}\bar{x} \tag{31}$$

and

$$\mathbb{P}(\mathrm{C}) = \bar{\epsilon}\, ax + \alpha a\bar{x} + \beta\bar{a}x + \gamma\bar{a}\bar{x}. \tag{32}$$

Hence,

$$\Delta_B = \frac{rx\bar{x}}{rx + \mu\bar{r}}\cdot\widetilde{\phi}_B, \tag{33}$$

with

$$\widetilde{\phi}_B := (\bar{\alpha} - \epsilon)a + (\beta - \gamma)\bar{a}. \qquad \square \tag{34}$$

Theorem 9. *Consider the Bayesian Network in Fig. 10 with the probability distribution defined in Eqs. (3), (5) and (8). Then (i)* $\Delta_C > 0$ *iff* $\bar{\alpha} > \epsilon$, *(ii)* $\Delta_C = 0$ *iff* $\bar{\alpha} = \epsilon$, *and (iii)* $\Delta_C < 0$ *iff* $\bar{\alpha} < \epsilon$.

Proof. Similarly, we obtain

$$\mathbb{P}(\mathrm{C}|\mathrm{A}) = \bar{\epsilon}\, x + \alpha\bar{x} \tag{35}$$

and

$$\mathbb{P}(\mathrm{A}, \mathrm{C}, \mathrm{Rep_X}) = a[\bar{\epsilon}\, rx + \mu\bar{r}(\bar{\epsilon}\, x + \alpha\bar{x})]. \tag{36}$$

Using Eq. (15), we finally obtain

$$\Delta_C = \frac{rx\bar{x}}{rx + \mu\bar{r}}\cdot\left(\bar{\alpha} - \epsilon\right), \tag{37}$$

from which the theorem follows. □

Theorem 10. *Consider the Bayesian Network in Fig. 12 with the probability distribution defined in Eqs. (3), (4) and (9). Then* $\Delta_A > 0$.

Proof. We calculate the new expressions for $P(\mathrm{Rep_X})$ and $P(\mathrm{A}, \mathrm{Rep_X})$ and obtain:

$$\mathbb{P}(\mathrm{Rep_X}) = (a + \beta\bar{a})rx + \mu\bar{r} \tag{38}$$

$$\mathbb{P}(\mathrm{A}, \mathrm{Rep_X}) = a(rx + \mu\bar{r}) \tag{39}$$

Hence, as $a + \beta\bar{a} < 1$,

$$\mathbb{P}\left(\mathrm{A}|\mathrm{Rep_X}\right) = \frac{rx + \mu\bar{r}}{(a + \beta\bar{a})rx + \mu\bar{r}}\cdot a > a, \tag{40}$$

from which the theorem follows. □

Theorem 11. *Consider the Bayesian Network in Fig. 12 with the probability distribution defined in Eqs. (3), (4) and (9). Then* $\Delta_B > 0$.

Proof. We calculate

$$\mathbb{P}(\mathrm{C}, \mathrm{Rep_X}) = (a + \beta\bar{a})rx + \mu\bar{r}\cdot P(\mathrm{C}) \tag{41}$$

and use Eq. (38) to obtain

$$\Delta_B = \frac{(a + \beta\bar{a})rx}{(a + \beta\bar{a})rx + \mu\bar{r}\cdot P(\mathrm{C})}\cdot P\left(\neg\,\mathrm{C}\right), \tag{42}$$

which is greater than zero. □

Theorem 12. *Consider the Bayesian Network in Fig. 12 with the probability distribution defined in Eqs. (3), (4) and (9). Then* $\Delta_C > 0$.

Proof. We calculate

$$\begin{aligned} \mathbb{P}(\mathrm{C}|\mathrm{A}) &= x + \alpha\bar{x} \\ \mathbb{P}(\mathrm{A}, \mathrm{C}, \mathrm{Rep_X}) &= a[rx + \mu\bar{r}(x + \alpha\bar{x})] \end{aligned} \tag{43}$$

and use Eq. (39) to obtain

$$\Delta_C = \frac{\bar{\alpha}rx\bar{x}}{rx + \mu\bar{r}(x + \alpha\bar{x})}, \tag{44}$$

which is greater than zero. □

Theorem 13. *Consider the Bayesian Network in Fig. 13 with the probability distribution defined in Eqs. (3), (4) and (10). Then* $\Delta_A > 0$.

Proof. We calculate the new expressions for $P(\mathrm{Rep_X})$ and $P(\mathrm{A}, \mathrm{Rep_X})$ and obtain:

$$\mathbb{P}(\mathrm{Rep_X}) = arx + \mu\bar{r} \tag{45}$$

$$\mathbb{P}(\mathrm{A}, \mathrm{Rep_X}) = a(rx + \mu\bar{r}) \tag{46}$$

Hence,

$$\mathbb{P}\left(\mathrm{A}|\mathrm{Rep_X}\right) = \frac{rx + \mu\bar{r}}{arx + \mu\bar{r}}\cdot a > a, \tag{47}$$

from which the theorem follows. □

Theorem 14. *Consider the Bayesian Network in Fig. 13 with the probability distribution defined in Eqs. (3), (4) and (10). Then* $\Delta_B > 0$.

Proof. We calculate

$$\mathbb{P}(\mathrm{C}, \mathrm{Rep_X}) = arx + \mu\bar{r}\cdot P(\mathrm{C}) \tag{48}$$

and use Eq. (38) to obtain

$$\Delta_B = \frac{arxP(\neg\,\mathrm{C})}{arx + \mu\bar{r}}, \tag{49}$$

which is greater than zero. □

Theorem 15. *Consider the Bayesian Network in Fig. 13 with the probability distribution defined in Eqs. (3), (4) and (10). Then* $\Delta_C > 0$.

Proof. We calculate

$$\begin{aligned} \mathbb{P}(\mathrm{C}|\mathrm{A}) &= x + \alpha\bar{x} \\ \mathbb{P}(\mathrm{A}, \mathrm{C}, \mathrm{Rep_X}) &= a[rx + \mu\bar{r}(x + \alpha\bar{x})] \end{aligned} \tag{50}$$

and use Eq. (39) to obtain

$$\Delta_C = \frac{\bar{\alpha}rx\bar{x}}{rx + \mu\bar{r}}, \tag{51}$$

which is greater than zero. □

Appendix C. The interpretation of "*X*"

The approach taken by the Modified Model is that it attempts to model the effect of learning an indicative conditional by saying as little as possible about what an indicative conditional amounts to. Even so, from a theoretical standpoint, the model must pair an interpretation of probability $\mathbb{P}$—whether $\mathbb{P}$ is the decision maker's personal probability distribution that parameterizes the graph or is the decision modeler's—with a consistent interpretation of the 5 random variables in the model. This note explores some issues to consider in pursuing that end.

According to the canonical theories of subjective probability due to Ramsey (1926), de Finetti (1937), Savage (1954), Anscombe and Aumann (1963), DeGroot (1970), an agent's partial beliefs are coherent if, and only if, they are representable by a probability function, and an agent's partial beliefs ought to be coherent to be rational. Why? According to de Finetti's notion of coherence, an agent's partial beliefs are the prices for some (finite) number of gambles that she deems fair. To assess the rationality of her partial beliefs, she is imagined to commit to either buy or sell any finite combination of those gambles at those prices. She exposes herself to the possibility of suffering a sure loss if and only if her partial beliefs fail to satisfy the axioms of (finitely additive) probability (de Finetti, 1937). He later introduced a second notion of coherence, constructed in terms of a strictly proper scoring rule, and a theorem to connect the two notions of coherence (de Finetti, 1974).

While de Finetti's coherence arguments are familiar, less appreciated is his allowance for incompletely specified information (de Finetti & Savage, 1962) and, as a result, permitting agents to have coherent commitments that nevertheless may have indeterminate consequences. Even if we insist on fully specifying information to eliminate logical ambiguity, the agent still may not know whether a particular state is logically consistent with that information, and several weakened notions of possibility are on offer (Walley, 1991, §2.1).

The point is that coherence depends on the states the agent considers possible, but different interpretations of 'possible' yield different effects on the strength and scope of the rationality norms associated with the theory (Wheeler, 2018). One notion of possibility will include all states that are logically consistent with the agent's available information, whereas another only those for which the agent has not determined are inconsistent with the available information. According to the former "epistemic possibility", there is no possibility of the agent being uncertain what the 9th digit of pi is, whereas according to the latter, "apparent possibility", an agent may be uncertain about what the 9th digit of pi. Suppose that you trust us and we tell you: '*if it rains, then the picnic is postponed.*' What are you presumed by your model to learn? More generally, what do you learn when a reliable witness tells you, 'If A, then B'? Worries mount whether these are coherent questions.

The approach that we take here interprets 'If A, then B' as an entailment relation between A and B. However, and crucially for our account, we deliberately leave unspecified both the structure and content of the binary relation between A and B (Kyburg, Teng, & Wheeler, 2007). Thus, in addition to an agent's partial belief in A and partial belief in B on the supposition that A, we shall also consider the agent's partial belief in $\mathrm{A} \vdash \mathrm{B}$, which in our model may be read either as 'A entails B' or 'If A, then B', and encoded by $X = \mathrm{X}$.

Suppose now that A and B are binary random variables, where the partial belief that A and the conditional partial belief that A given B are representable by $\mathbb{P}(\mathrm{A})$ and $\mathbb{P}(\mathrm{B}|\mathrm{A})$, respectively, for a finitely additive probability $\mathbb{P}(\cdot)$. Unlike partial belief that A entails B, which we will return to in a moment, conditional judgments representable by a probability mass function carry a number of constraints. First, unconditional and conditional commitments are *synchronic* in the sense that both are assessed for coherence at a specific moment in time. Although this is not a complication for the partial belief that B, i.e., $\mathbb{P}(\mathrm{B})$, the criteria of coherence does constrain the interpretation of the partial belief that B conditional on A, i.e., $\mathbb{P}(\mathrm{B}|\mathrm{A})$. There are two subtly different options for interpreting the conditional probability $\mathbb{P}(\mathrm{B}|\mathrm{A})$:

Called-off interpretation: $\mathbb{P}(\mathrm{B}|\mathrm{A})$: expresses the agent's current unconditional partial belief that B contingent on the occurrence of A.

Example: $\mathbb{P}(\mathrm{B}|\mathrm{A}) = p$ elicited from you at time t, according to de Finetti, expresses your commitment at t to accept all contracts of value α concerning $B = \mathrm{B}$ such that $\mathbf{1}_B\alpha(\mathbf{1}_A - p)$ as *fair*, which are called-off (i.e., yield you 0) if $B = \neg\,\mathrm{B}$, that is, if B does not occur.[48]

Hypothetical interpretation: $\mathbb{P}(\mathrm{B}|\mathrm{A})$ expresses the agent's current unconditional partial belief that $B = \mathrm{B}$ contingent on the hypothetical supposition that $A = \mathrm{A}$.

Example: Savage's Savage (1954) distinction between a basic decision problem and a derived decision problem is a contrast between your current preference over pairs of acts versus your current preference over the same pairs of acts on the hypothetical supposition that the event $A = \mathrm{A}$ occurs. Savage shows that the expected value of a gamble in a derived decision problem cannot be less than, and is often greater than, the expected value of the corresponding gamble in a basic decision problem.[49]

The interpretation of $\mathrm{A} \vdash \mathrm{B}$, by contrast, is far less constrained. Rather than commit to a particular theory of "If-then" sentences, our model makes hardly any theoretical commitments to "If-then" sentences and explicitly leaves out the extra content and structure that one might anticipate is necessary to interpret the model. Instead, we treat the assertion that $\mathrm{A} \vdash \mathrm{B}$ as the positive value of another binary random variable that we label, X. Specifically, $X = \mathrm{X}$ codes '*The indicative conditional* "If A, then B" *holds*,' and $X = \neg\,\mathrm{X}$ codes '*The indicative conditional* "If A, then B" *does not hold.*'.

There are two remaining random variables in the model, REP_X and REL. $REP_X = \mathrm{REP_X}$ expresses '*The information source reports* $X = \mathrm{X}$' and $REP_X = \neg\,\mathrm{Rep_X}$ expresses '*The information source does not report* $X = \mathrm{X}$.' $REL = \mathrm{Rel}$ expresses '*The information source is reliable*' and $REL = \neg\,\mathrm{Rel}$ expresses '*The information source is not reliable.*'.

[48] Note that $\mathbf{1}_B$ is the "indicator function" applied to B, returning 1 if B occurs and 0 otherwise.

[49] For a discussion of the hypothetical and called-off interpretation and their roles in decision making, see (Pedersen & Wheeler, 2015).

References

Adams, E. W. (1975). *The logic of conditionals.* Dordrecht: D. Reidel.

Anscombe, F. J., & Aumann, R. J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics, 34*, 199–205.

Asher, N., & Lascarides, A. (2003). *Logics of conversation.* Cambridge: Cambridge University Press.

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

Baratgin, J., & Politzer, G. (2010). Updating: A psychologically basic situation of probability revision. *Thinking & Reasoning, 16*, 245–287.

Baratgin, J., Politzer, G., & Over, D. E. (2013). Uncertainty and the de Finetti tables. *Thinking & Reasoning, 19*(3), 308–328.

Baratgin, J., Politzer, G., Over, D. E., & Takahashi, T. (2018). The psychology of uncertainty and three-valued truth tables. *Frontiers in Psychology, 9*(1479), https://doi.org/10.3389/fpsyg.2018.01479.

Barrouillet, P., & Gauffroy, C. (2015). Probability in reasoning: A developmental test on conditionals. *Cognition, 137*, 22–39.

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

Berwick, R. C., Pietroski, P., Yankama, B., & Chomsky, N. (2011). Poverty of the stimulus revisited. *Cognitive Science, 35*(7), 1207–1242.

Bovens, L., Fitelson, B., Hartmann, S., & Snyder, J. (2002). Too odd (not) to be true? A reply to Olsson. *The British Journal for the Philosophy of Science, 53*(4), 539–563.

Bovens, L., & Hartmann, S. (2002). Bayesian networks and the problem of unreliable instruments. *Philosophy of Science, 69*(1), 29–72. https://doi.org/10.1086/338940 Retrieved 2015–01-28TZ, from http://www.jstor.org/stable/10.1086/338940.

Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology.* Oxford: Oxford University Press.

Bradley, R. (2005). Radical probabilism and bayesian conditioning. *Philosophy of Science, 72*(2), 342–364.

Briñol, P., & Petty, R. E. (2009). Source factors in persuasion: A self-validation approach. *European Review of Social Psychology, 20*(1), 49–96. https://doi.org/10.1080/10463280802643640 Retrieved 2015–09-09TZ.

Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality.* Cambridge, MA: The MIT Press.

Chomsky, N. (1988). *Language and problems of knowledge: The managua lectures, vol. 16*. MIT Press.

Collins, P. J. (2017). *Rationality, pragmatics, and sources* (Unpublished doctoral dissertation)Birkbeck: University of London.

Collins, P. J., Hahn, U., von Gerber, Y., & Olsson, E. J. (2018). The bi-directional relationship between source characteristics and message content. *Frontiers in Psychology, 9*, 18.

Cruz, N., & Oberauer, K. (2014). Comparing the meanings of "if" and "all". *Memory & Cognition, 42*(8), 1345–1356.

Csiszár, I. (1967). Information–type measures of difference of probability distributions i–ii, mta ii. *Oszt. Közl, 17*, 123–149.

Declerck, R., & Reed, S. (2001). *Conditionals: A comprehensive empirical analysis.* Berlin: Walter de Gruyter (Google-Books-ID: RtyewWWe1LAC).

de Finetti, B. (1936/1995). The logic of probability. *Philosophical Studies, 77*, 181–190.

de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré, 7*(1), 1–68 (Trans. Henry E. Kyburg, Jr., Foresight: Its logical laws, its subjective sources. In H.E. Kyburg, Jr. and H. Smokler (Eds.), Studies in Subjective Probability. New York: John Wiley and Sons, 1964).

de Finetti, B. (1974). *Theory of probability: A critical introductory treatment, Vol. 1 and 2*. Wiley.

de Finetti, B., & Savage, L. J. (1962). Sul modo di scegliere le probabilità iniziali. *Biblioteca del Metron, Serie C, 1*, 81–154.

DeGroot, M. H. (1970). *Optimal statistical decisions.* New York: McGraw-Hill.

Diaconis, P., & Zabell, S. (1982). Updating subjective probability. *Journal of the American Statistical Association, 77*, 822–830.

Dietz, R., & Douven, I. (2010). Ramsey's test, Adams' thesis, and left-nested conditionals. *The Review of Symbolic Logic, 3*(3), 467–484.

Douven, I. (2012). Learning conditional information. *Mind & Language, 27*(3), 239–263.

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

Douven, I. (2016). On de Finetti on iterated conditionals. In C. Beierle, G. Brewka, & M. Thimm (Eds.). *Computational models of rationality: Essays dedicated to Gabriele Kern-Isberner on the occasion of her 60th birthday*. College Publications.

Douven, I., & Dietz, R. (2011). A puzzle about Stalnaker's hypothesis. *Topoi, 30*(1), 31–37.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2018). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology, 101*, 50–81. https://doi.org/10.1016/j.cogpsych.2017.09.002.

Douven, I., & Romeijn, J. W. (2011). A new resolution of the Judy Benjamin problem. *Mind, 120*(479), 637–670.

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

Dubins, L. E., et al. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability, 3*(1), 89–99.

Edgington, D. (1995a). On conditionals. *Mind, 104*(414), 235–329. https://doi.org/10.1093/mind/104.414.235 Retrieved 2017–02-14TZ, from https://academic.oup.com/mind/article-lookup/doi/10.1093/mind/104.414.235.

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

Elder, C. H., & Jaszczolt, K. M. (2016). Towards a pragmatic category of conditionals. *Journal of Pragmatics, 98*, 36–53. https://doi.org/10.1016/j.pragma.2016.04.013 Retrieved 2017–04-10TZ, from http://www.sciencedirect.com/science/article/pii/S0378216616301394.

Elqayam, S., & Evans, J. S. B. T. (2011). Subtracting ought from is: Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences, 34*, 233–248. https://doi.org/10.1080/13546783.2013.834268.

Eva, B., & Hartmann, S. (2019). *On the origins of old evidence*. (To appear in Australasian Journal of Philosophy, preprint available under http://philsci-archive.pitt.edu/15778/).

Eva, B., Hartmann, S., & Rad, S.R. (2019). *Learning from conditionals*. (To appear in Mind, preprint available under http://philsci-archive.pitt.edu/15835/).

Eva, B., & Hartmann, S. (2018). Bayesian argumentation and the value of logical validity. *Psychological Review, 125*(5), 806–821.

Evans, J.S.B.T., Newstead, S.E., & Byrne, R.M.J. (1993). *Human reasoning: The psychology of deduction*. Hove, England: Lawrence Erlbaum Associates. (Google-Books-ID: iFMhZ4dl1KcC).

Eva, B., Hartmann, S., & Singmann, H. (2019). *A new probabilistic explanation of the modus ponens-modus tollens asymmetry*. Montreal, QB: Cognitive Science Society289–294.

Evans, J. S. B. T., Handley, S. J., Neilens, H., & Over, D. E. (2007). Thinking about conditionals: A study of individual differences. *Memory & Cognition, 35*(7), 1772–1784. https://doi.org/10.3758/BF03193509 Retrieved 2017–02-14TZ, from http://www.springerlink.com/index/10.3758/BF03193509.

Evans, J. S. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(2), 321–335. https://doi.org/10.1037/0278-7393.29.2.321 Retrieved 2017–02-14TZ, from http://doi.apa.org/getdoi.cfm?doi=10.1037/0278-7393.29.2.321.

Evans, J. S. B. T., Neilens, H., Handley, S. J., & Over, D. E. (2008). When can we say 'if'? *Cognition, 108*(1), 100–116. https://doi.org/10.1016/j.cognition.2008.02.001 Retrieved 2015–02-05TZ, from http://www.sciencedirect.com/science/article/pii/S0010027708000310.

Franke, M. & Jäger, G. (2016). Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics. *Zeitschrift für Sprachwissenschaft* 35(1), Retrieved 2016–10-11TZ, from http://www.degruyter.com/view/j/zfsw.2016.35.issue-1/zfs-2016-0002/zfs-2016-0002.xml. doi:10.1515/zfs-2016-0002.

Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts toward the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(3), 635–648. https://doi.org/10.1037/a0022329 Retrieved 2017–02-14TZ, from http://doi.apa.org/getdoi.cfm?doi=10.1037/a0022329.

Garber, D. (1983). Old evidence and logical omniscience in bayesian confirmation theory. In J. Earman (Ed.). *Testing scientific theories* (pp. 99–131). Minneapolis: The University of Minnesota Press.

Gärdenfors, P. (1988). *Knowledge in flux: Modelling dynamics of epistemic states.* Cambridge, MA: The MIT Press.

Goodman, N. D., & Lassiter, D. (2015). Probabilistic semantics and pragmatics uncertainty in language and thought. In S. Lappin, & C. Fox (Eds.). *The handbook of contemporary semantic theory* (pp. 655–686). John Wiley & Sons, Ltd.. https://doi.org/10.1002/9781118882139.ch21.

Goodwin, G. P. (2014). Is the basic conditional probabilistic? *Journal of Experimental Psychology: General, 143*(3), 1214–1241. https://doi.org/10.1037/a0034232 Retrieved 2017–02-14TZ, from http://doi.apa.org/getdoi.cfm?doi=10.1037/a0034232.

Grice, H. P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Günther, M. (2017). Learning conditional information by Jeffrey imaging on Stalnaker conditionals. *Journal of Philosophical Logic*. https://doi.org/10.1007/s10992-017-9452-z.

Hahn, U., Harris, A. J., & Corner, A. (2009). Argument content and argument source: An exploration. *Informal Logic, 29*(4), 337–367.

Hahn, U., Harris, A. J. L., & Corner, A. (2016). Public reception of climate science: coherence, reliability, and independence. *Topics in Cognitive Science, 8*(1), 180–195. https://doi.org/10.1111/tops.12173 Retrieved 2017–02-25TZ, from http://doi.wiley.com/10.1111/tops.12173.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review, 114*(3), 704–732. https://doi.org/10.1037/0033-295X.114.3.704.

Haigh, M., Stewart, A.J., & Connell, L. (2013). Reasoning as we read: Establishing the probability of causal conditionals. *Memory* & *Cognition* 41(1), 152–158. Retrieved 2016–10-11TZ, from http://link.springer.com/article/10.3758/s13421-012-0250-0. doi:10.3758/s13421-012-0250-0.

Hahn, U., Oaksford, M., & Harris, A. (2012). Testimony and argument: A Bayesian perspective. In F. Zenker (Ed.). *Bayesian argumentation* (pp. 15–38). New York, NY: Springer.

Harper, W. L. (1976). Ramsey test conditionals and interated belief change. In W. L. Harper, & C. A. Hooker (Eds.). *Foundations of probability theory, statistical inference, and statistical theories of science* (pp. 117–135). Dordrecht: D. Reidel.

Harris, A. J. L., Hahn, U., Madsen, J. K., & Hsu, A. S. (2016). The appeal to expert opinion: Quantitative support for a Bayesian network approach. *Cognitive Science, 40*(6), 1496–1533. https://doi.org/10.1111/cogs.12276 Retrieved 2016–10-11TZ, from http://doi.wiley.com/10.1111/cogs.

Hartmann, S., & Fitelson, B. (2015). A new Garber-style solution to the problem of old evidence. *Philosophy of Science, 82*(4), 712–717.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin, 118*(2), 248–271. https://doi.org/10.1037/0033-2909.118.2.248.

Hinterecker, T., Knauff, M., & Johnson-Laird, P. N. (2016). Modality, probability, and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(10), 1606–1620. https://doi.org/10.1037/xlm0000255.

Jackson, F. (1979). On assertion and indicative conditionals. *The Philosophical Review, 88*(4), 565–589.

Jarvstad, A., & Hahn, U. (2011). Source reliability and the conjunction fallacy. *Cognitive Science, 35*(4), 682–711. https://doi.org/10.1111/j.1551-6709.2011.01170.x Retrieved 2015–01-28TZ, from http://onlinelibrary.wiley.com/doi/10.1111/j.1551-6709.2011.01170.x/abstract.

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

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

Johnson-Laird, P., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences, 19*(4), 201–214. https://doi.org/10.1016/j.tics.2015.02.006 Retrieved 2016–12-07TZ, from http://linkinghub.elsevier.com/retrieve/pii/S1364661315000303.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*(2), 136–153.

Kern-Isberner, G. (2001). *Conditionals in nonmonotonic reasoning and belief revision.* Berlin/Heidelberg: Springer.

Khemlani, S. S., Byrne, R. M. J., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive Science, 42*(6), 1887–1924. https://doi.org/10.1111/cogs.12634.

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

Kratzer, A. (1979). Conditional necessity and possibility. In R. Bäuerle, U. Egli, & A. von Stechow (Eds.). *Semantics from different points of view*. Berlin/Heidelberg: Springer.

Kratzer, A. (1986). Conditionals. *Chicago Linguistic Society, 22*(2), 1–15.

Korta, K., & Perry, J. (2015). Pragmatics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2015 ed.). Retrieved 2017–01-09, from https://plato.stanford.edu/archives/win2015/entries/pragmatics/.

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

Kyburg, H. E., Jr., Teng, C. M., & Wheeler, G. (2007). Conditionals and consequences. *Journal of Applied Logic, 5*(4), 638–650.

Landes, J., Osimani, B., & Poellinger, R. (2018). Epistemology of causal inference in pharmacology. *European Journal for Philosophy of Science, 8*(1), 3–49.

Leahy, B., Rafetseder, E., & Perner, J. (2014). Basic conditional reasoning: How children mimic counterfactual reasoning. *Studia Logica,* 1–18.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalised conversational implicature.* Cambridge, MA: The MIT Press.

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

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *The Philosophical Review, 85*(3), 297–315.

Lucas, C. G., & Kemp, C. (2015). An improved probabilistic account of counterfactual reasoning. *Psychological Review, 122*(4), 700–734.

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition, 46*(1), 53–85.

Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology, 29*, 87–109.

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology, 49*(3), https://doi.org/10.1037/a0029500 Retrieved 2015–07-28TZ, from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3810952/.

Krzyżanowska, K., Collins, P.J., & Hahn, U. (2019). True clauses and false connections, submitted for publication.

Mirabile, P., & Douven, I. (in press). Abductive conditionals as a test case for inferentialism. *Cognition*. doi:10.1016/j.cognition.2020.104232.

Newell, A., & Simon, H. A. (1961). Computer simulation of human thinking. *Science, 134*(3495), 2011–2017.

Oaksford, M. & Chater, N. (2019). New paradigms in the psychology of reasoning. *Annual Review of Psychology, 71*.

Oaksford, M., & Chater, N. (1991). Against logicist cognitive science. *Mind & Language, 6*(1), 1–38.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*(4), 608–631. https://doi.org/10.1037/0033-295X.101.4.608.

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

Oaksford, M., & Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Thinking & Reasoning, 19*(3–4), 346–379. https://doi.org/10.1080/13546783.2013.808163 Retrieved 2015–05-24TZ.

Oaksford, M., Over, D., & Cruz, N. (2019). Paradigms, possibilities, and probabilities: Comment on Hinterecker, Knauff, and Johnson-Laird (2016). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(2), 288–297.

Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 680–693. https://doi.org/10.1037/0278-7393.29.4.680 Retrieved 2017–02-14TZ, from http://doi.apa.org/getdoi.cfm?doi=10.1037/0278-7393.29.4.680.

Ohm, E., & Thompson, V. A. (2006). Conditional probability and pragmatic conditionals: Dissociating truth and effectiveness. *Thinking & Reasoning, 12*(3), 257–280.

Over, D. E. (2016). The paradigm shift in the psychology of reasoning. In L. Macchi, M. Bagassi, & R. Viale (Eds.). *Cognitive unconscious and human rationality* (pp. 79–98). Cambridge, MA: MIT Press.

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

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

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

Over, D. E., Hadjichristidis, C., Evans, J. S. B., Handley, S. J., & Sloman, S. A. (2007b). The probability of causal conditionals. *Cognitive Psychology, 54*(1), 62–97. https://doi.org/10.1016/j.cogpsych.2006.05.002 Retrieved 2017–02-14TZ, from http://linkinghub.elsevier.com/retrieve/pii/S0010028506000478.

Pagin, P. (2016). Assertion. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2016 ed.). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/win2016/entries/assertion/.
Pearl, J. (2000). *Causality: Models, reasoning, and inference.* Cambridge University Press.
Pearl, J. (2013). Structural counterfactuals: A brief introduction. *Cognitive Science, 37*(6), 977–985.
Perner, J., & Rafetseder, E. (2011). Counterfactual and other forms of conditional reasoning: Children lost in the nearest possible world. In C. Hoerl, T. McCormack, & S. Beck (Eds.). *Understanding counterfactuals, understanding causation. Issues in philosophy and psychology*. Oxford University Press.
Politzer, G., Over, D. E., & Baratgin, J. (2010). Betting on conditionals. *Thinking & Reasoning, 16*(3), 172–197. https://doi.org/10.1080/13546783.2010.504581 Retrieved 2017–02-14TZ, from http://www.tandfonline.com/doi/abs/10.1080/13546783.2010.504581.
Popper, Karl, & Miller, David (1983). A proof of the impossibility of inductive logic. *Nature, 302*, 687–688. https://doi.org/10.1038/302687a0.
Popper, K. R. (1959). The propensity interpretation of probability. *The British Journal for the Philosophy of Science, 10*(37), 25–42.
Quelhas, A. C., Johnson-Laird, P. N., & Juhos, C. (2010). The modulation of conditional assertions and its effects on reasoning. *The Quarterly Journal of Experimental Psychology, 63*(9), 1716–1739.
Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of nearest possible world. *Child Development, 81*(1), 376–389.
Pedersen, A. P., & Wheeler, G. (2015). Dilation, disintegrations, and delayed decisions. In *Proceedings of the 9th symposium on imprecise probabilities and their applications (ISIPTA)* (pp. 227–236). Pescara, Italy.
Ramsey, F. P. (1926). Truth and probability. In H. E. Kyburg & H. E. Smokler (Eds.), *Studies in subjective probability* (Second (1980) ed., pp. 23–52). Huntington, New York: Robert E. Krieger Publishing Company.
Ramsey, F. P. (1929/1990). General propositions and causality. In D. H. Mellor (Ed.). *Philosophical papers* (pp. 145–163). Cambridge: Cambridge University Press.
Recanati, F. (2010). *Truth-conditional pragmatics.* Oxford: Oxford University Press.
Rényi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Hungarica, 6*(3–4), 285–335.
Rieger, A. (2013). Conditionals are material: the positive arguments. *Synthese, 190*(15), 3161–3174. https://doi.org/10.1007/s11229-012-0134-7.
Rips, L. J., & Marcus, S. L. (1977). Suppositions and the analysis of conditional sentences. In M. A. Just, & P. A. Carpenter (Eds.). *Cognitive processes in comprehension* (pp. 185–220). Hillsdale, NJ: Lawrence Erlbaum.
Sanfilippo, G., Gilio, A., Over, D. E., & Pfeifer, N. (2020). Probabilities of conditionals and previsions of iterated conditionals. *International Journal of Approximate Reasoning, 121*, 150–173. https://doi.org/10.1016/j.ijar.2020.03.001.
Sanfilippo, G., Pfeifer, N., Over, D. E., & Gilio, A. (2018). Probabilistic inferences from conjoined to iterated conditionals. *International Journal of Approximate Reasoning, 93*, 103–118.
Savage, L. J. (1954). *Foundations of statistics.* New York: Wiley.
Schroyens, W. (2010). A meta-analytic review of thinking about what is true, possible, and irrelevant in reasoning from or reasoning about conditional propositions. *European Journal of Cognitive Psychology, 22*(6), 897–921. https://doi.org/10.1080/09541440902928915 Retrieved 2017–05-26TZ.
Schwan, B., & Stern, R. (2007). A causal understanding of when and when not to jeffrey conditionalize. *Philosopher's Imprint, 17*(8), 1–21.
Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation.* Mahwah, New Jersey: Lawrence Erlbaum Associates (Google-Books-ID: yGUAAwAAQBAJ).
Skovgaard-Olsen, N., Collins, P., Krzyżanowska, K., Hahn, U., & Klauer, K. C. (2019). Cancellation, negation, and rejection. *Cognitive Psychology, 108*, 42–71.
Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016a). Relevance and reason relations. *Cognitive Science, 1202–1215*. https://doi.org/10.1111/cogs.12462 Retrieved 2017–03-06TZ, from http://onlinelibrary.wiley.com/doi/10.1111/cogs.12462/abstract.
Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016b). The relevance effect and conditionals. *Cognition, 150*, 26–36.
Sobel, J. H. (2004). Pascalian wagers. In *Logic and theism: Arguments for and against beliefs in god* (pp. 499-537). Cambridge University Press.
Sprenger, J., & Hartmann, S. (2019). *Bayesian philosophy of science.* Oxford: Oxford University Press.
Stalnaker, R. C. (1968). A theory of conditionals. In N. Rescher (Ed.). *Studies in logical theory* (pp. 98–112). Oxford: Blackwell.
Stalnaker, R. C. (1970). Probability and conditionals. *Philosophy of Science, 37*, 64–80.
Stalnaker, R. C. (1975). Indicative conditionals. *Philosophia, 5*(3), 269–286.
Stalnaker, R. C. (1976). Stalnaker to van Fraassen. In W. L. Harper, & C. A. Hooker (Vol. Eds.), *Foundations of probability theory, statistical inference, and statistical theories of science: Vol. 1*, (pp. 302–306). Dordrecht: D. Reidel.
Stalnaker, R. C. (1984). *Inquiry.* The MIT Press.
Steels, L. (1985). Second generation expert systems. *Future Generation Computer Systems, 1*(4), 213–221.
Stenning, K., & Lambalgen, M.v. (2008). *Human reasoning and cognitive science.* A Bradford Book.
Stern, R., & Hartmann, S. (2018). Two sides of modus ponens. *The Journal of Philosophy 115*(11), 605–621.
Stevenson, R. J., & Over, D. E. (2001). Reasoning from uncertain premises: Effects of expertise and conversational context. *Thinking & Reasoning, 7*(4), 367–390. https://doi.org/10.1080/13546780143000080 Retrieved 2016–04-08TZ, from http://www.tandfonline.com/doi/abs/10.1080/13546780143000080.
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,* 1154–1170.
van Fraassen, B. C. (1981). A problem for relative information minimizers in probability kinematics. *The British Journal for the Philosophy of Science, 32*(4), 375–379.
Walley, P. (1991). *Statistical reasoning with imprecise probabilities.* London: Chapman and Hall.
Wheeler, G. (2018). Bounded rationality. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2018 ed.). Metaphysics Research Lab, Stanford University; Retrieved from https://plato.stanford.edu/entries/bounded-rationality/.
Wheeler, G., & Scheines, R. (2013). Coherence, confirmation, and causation. *Mind, 122*(435), 135–170.
Zhao, J., & Osherson, D. (2014). Category-based updating. *Thinking & Reasoning, 20*(1), 1–15. https://doi.org/10.1080/13546783.2013.795916.