Formal Epistemology and New Paradigm Psychology of Reasoning

Pfeifer, Niki; Douven, Igor

Published in:
Review of Philosophy and Psychology

DOI:
10.1007/s13164-013-0165-0

Rev.Phil.Psych. (2014) 5:199–221
DOI 10.1007/s13164-013-0165-0

Formal Epistemology and the New Paradigm Psychology of Reasoning

Niki Pfeifer · Igor Douven

Published online: 26 September 2013

Abstract This position paper advocates combining formal epistemology and the new paradigm psychology of reasoning in the studies of conditionals and reasoning with uncertainty. The new paradigm psychology of reasoning is characterized by the use of probability theory as a rationality framework instead of classical logic, used by more traditional approaches to the psychology of reasoning. This paper presents a new interdisciplinary research program which involves both formal and experimental work. To illustrate the program, the paper discusses recent work on the paradoxes of the material conditional, nonmonotonic reasoning, and Adams' Thesis. It also identifies the issue of updating on conditionals as an area which seems to call for a combined formal and empirical approach.

This paper highlights the emergence of a new research program concerning uncertain reasoning. The program—henceforth simply referred to as "the new program"—is situated at the intersection of formal epistemology and experimental psychology of reasoning, and combines formal and empirical methods to address issues that call for the participation of both philosophers and psychologists.

Probabilistic approaches to conditionals have been popular in philosophy for some decades now (e.g., Adams 1975; Bennett 2003; Douven 2008). Current contributions on conditionals by philosophers working in the probabilistic paradigm mostly fall within the relatively young discipline of formal epistemology. The traditional paradigm psychology of reasoning used logic as the rationality framework. Over the

N. Pfeifer (✉)
Munich Center for Mathematical Philosophy, LMU Munich, Munich, Germany
e-mail: niki.pfeifer@lrz.uni-muenchen.de

I. Douven
Faculty of Philosophy, University of Groningen, Groningen, The Netherlands
e-mail: i.e.j.douven@rug.nl

past ten years, however, psychologists of reasoning have also become increasingly interested in studying conditionals from a probabilistic perspective (e.g., Oaksford and Chater 2009; Over et al. 2007; Pfeifer 2012; Pfeifer and Kleiter 2009). This shift from logic to probability theory as the rationality framework paved the way for the "new paradigm psychology of reasoning" (Evans 2012; Over 2009). We note that Oaksford and Chater have argued since the early 1990s against classical logic as the correct normative standard of reference for empirical reasoning research in favor of nonmonotonic reasoning (Oaksford and Chater 1991). Later, they began to systematically apply probabilistic models to classical deductive reasoning tasks, like their probabilistic analysis of the well-known Wason selection task (Oaksford and Chater 1994) and their probability heuristics model of categorical syllogisms (Chater and Oaksford 1999). However, the new paradigm psychology of reasoning only began to properly replace the traditional (logic-based) paradigm psychology about ten years ago: this probabilist turn emerged after the use of a much wider range of reasoning tasks like the probabilistic truth table tasks (Evans 2012).

This is not to suggest that formal epistemologists and psychologists studying conditionals have been addressing the same questions. For the most part, this is not true. For instance, formal epistemologists have been mainly concerned with questions about the truth conditions of conditionals, their assertability and acceptability conditions, and various notions of validity pertaining to arguments involving conditionals. Psychologists of reasoning, by contrast, have focused on how people interpret and process conditionals, and which inferences they are willing to make on the basis of conditional (and possibly other) premises.

Still, there have also been some clear points of contact between the two research communities. An example in this connection is the joint work by Jonathan Evans, an experimental psychologist, and David Over, who was trained as a formally-oriented philosopher before turning to psychology. The work of Evans and Over on the probabilities of conditionals (Evans and Over 2004), which is both philosophically informed and supported by important empirical results, has done much to popularize probabilistic approaches to conditionals in psychology but has also attracted attention from philosophers (Douven and Dietz 2011; Edgington 2003). And, to mention some contributions of our own, the first author (who was trained in psychology as well as in philosophy) has investigated paradoxes of the material conditional (Pfeifer 2013a; Pfeifer and Kleiter 2011), nonmonotonic reasoning (Pfeifer and Kleiter 2005, 2010), and Aristotle's Thesis (Pfeifer 2012) within a probability logical framework (see Sections 3 and 4); the second author has collaborated on experimental work on Adams' Thesis, which relates the assertability / acceptability of a conditional to the probability of its consequent given its antecedent (see Section 5).

We are convinced that there is something to be gained for both philosophers and psychologists from these double-sided and typically collaborative approaches. Where experimental data may help to confirm (or refute, as the case may be) philosophical theories, and may stimulate and guide the development of novel theories, formal epistemology provides tools to make psychological theorizing and hypothesis-building precise, and may in addition suggest rationality norms for evaluating human reasoning. Insofar as philosophical accounts can be shown to have empirical content—and formal work may be required to work out exactly what that content is—experimental data will help us to choose among such accounts in the same way in which they help us to choose among theories in the empirical sciences.

While, in our view, the core of the new program is situated at the intersection of formal epistemology and the psychology of reasoning, it is also closely related to the cognitive science of reasoning. Many of the topics that cognitive scientists work on are naturally seen as falling into the domain of the new program as well. Key examples are probabilistic approaches to causal reasoning (e.g., Sloman 2005; Waldmann and Hagmayer 2005), causal learning (e.g., Kemp et al. 2010), argumentation (e.g., Hahn and Oaksford 2006), and coherence (e.g., Harris and Hahn 2009). Nevertheless, we believe that there is at least a difference in emphasis between the cognitive science of reasoning and the new program, in that research in the former area is mainly focussed on modelling human inferential practices in an empirically adequate way, sometimes even explicitly dodging philosophical issues (see, e.g., Harris and Hahn 2009, Footnote 1, p. 1369), while work belonging to the new program typically has among its main aims helping answer philosophical questions.

There is a similar difference between the research goals of the psychology of reasoning and those of the new program. The main focus of the cognitive psychology of reasoning is on how people form representations of the premises and the conclusion and how they manipulate representations in order to draw inferences. Information about how people reason and about the computational modelling of human inference provides helpful theoretical and methodological input for the new program. But, again, this program focuses on philosophical problems, mostly ones that have arisen in the area of formal epistemology; its goal is not to model or to describe how people reason. In short, the main goal of the new program is to address—ideally, to solve—philosophical problems by combining formal and empirical work, and to this extent it is truly deserving of the epithet "new".

It is worth stressing that the new program is not about replacing conceptual analysis with empirical work. We believe that conceptual analysis has more than proven its value in finely crafted philosophical work throughout the past century and even longer. We also believe that, for some classes of philosophical problems, conceptual analysis is not only necessary but also sufficient for obtaining satisfactory solutions.[1] Analyses of the interdependence of rationality axioms, for example, do not seem to require any empirical work, nor does a welter of other technical work of the kind carried out in many branches of analytic philosophy. We also believe, however, that

[1]Thus, we do not accept Kauppinen's *negative experimentalism thesis*—which allegedly is accepted by some experimental philosophers—according to which "armchair reflection and informal dialogue are *not* reliable sources of evidence for (philosophically relevant) claims about folk concepts" (Kauppinen 2007, p. 97).

the solution of other classes of philosophical problems may profit substantially from empirical work. In this paper, we will list a number of examples of philosophical problems where the interaction of formal and empirical work is at least fruitful if not essential.

Specifically, we will highlight a number of research areas where recent progress seems to have depended on philosophers and psychologists joining forces, or where future progress seems to call for such joint ventures. Thematically, these research areas are related in that they all concern, in one way or another, uncertain or probabilistic reasoning. Methodologically, they are related in that they call for the use of both formal and empirical tools. That, in our opinion, justifies thinking of these areas as constituting a new research program in its own right. But first we describe some recent and not so recent developments in philosophy and psychology that paved the way for the new research program.

1 The Naturalistic Turn in Philosophy

Philosophy has traditionally involved a good deal of speculation, but then again, so has science. However, whilst the results of scientific speculation are ultimately to confront the data, and will be rejected in case of conflict with those data, there is the danger that philosophical speculation may go uncontrolled and can thereby lead to results that may be hard to identify as constituting progress in any way. Twentieth-century philosophy showed a growing awareness of this risk, but opinions have diverged on what measures are to be taken to make the endeavor less risky.

The logical empiricists' insistence that philosophers aim for the same rigor in their reasoning as mathematicians and logicians do in theirs made the philosophical profession appear more scientific in many ways. However, it did little to make assurances that philosophers were not really engaging in some kind of philosophical fiction writing, however formally dressed up.[2]

It was certainly a decisive step forward when, in the 1960s, post-positivist authors like Quine, Kuhn, Feyerabend, Hanson, and Laudan began to propagate the idea that, rather than trying (merely) to emulate the rigor of scientific inquiry, philosophers ought to start paying more attention to science itself, both to scientific theories and to the actual practice of science. Famously, Quine (1969) noted that psychological findings about perception, memory, trust, and so on, are of immediate importance to epistemologists, and that if one's epistemology is at odds with the best scientific theories concerning, for instance, vision, then that counts heavily against that epistemology (even if it does not refute it, given that our best scientific theories may still be wrong). And Kuhn, Laudan, and other historically-oriented philosophers of science argued that philosophical accounts of, for example, the evolution of science were to

[2] As an anonymous referee noted, the logical empiricists may have believed their verification theory of meaning, according to which a sentence is meaningful only if it is verifiable, to eliminate the possibility that philosophers were really writing fiction. But that theory was soon found to be untenable, if only because it does not live up to its own standards of meaningfulness.

be tested against the historical track record of science, and to be repudiated if they made the wrong predictions about what is to be found in that track record.[3]

In some quarters, experimental philosophy has been touted as a further step in the right direction. Philosophers are no longer just aiming to be as rigorous in their argumentation as scientists are, nor are they satisfied any longer with simply attending to scientific findings. Rather, philosophers have now themselves become scientists: the birth of the philosopher–scientist, or—should we say—the return of the natural philosopher familiar from the days of Galileo, Newton, and their likes! What this means, more specifically, is that philosophers have started addressing what appear to be unmistakably philosophical questions—such as questions about moral responsibility and blameworthiness, about free will, and about the nature of knowledge—with the help of unmistakably empirical research methods.

Does this recent development really constitute as much of a leap forward as many of its proponents like to believe? Since experimental philosophy is still in its infancy, it would be premature and also unfair to try to give a general verdict on the approach. What already *can* be said, however, is that some of the work that has come out of this approach is clearly valuable. Over the past decades, many have come to often accept blanket appeals to intuition as providing evidence for philosophical theses. Indeed, sometimes the bulk of the work that an author invests in defending a position goes into arguing that the position is in line with what we say intuitively / pre-reflectively / pre-analytically / pre-theoretically, or with how the folk think about certain matters. For instance, a number of epistemologists concerned with defending certain contextualist views of knowledge have proceeded mainly in this manner. According to these views, people's willingness to attribute knowledge of a given proposition to a subject is sensitive to what is at stake for the subject in being right about that proposition. The epistemologists in question typically argue for this claim by presenting their readers with different types of fictional vignettes about a person who has a certain amount of evidence for a proposition and also certain stakes in being right about the proposition. These stories are supposed to elicit different intuitive verdicts in the readers as to whether or not the person in the story knows the given proposition. Of course, what the authors then propose as *the* intuitive verdicts about the stories are simply their own, even though many readers seem to have agreed. A methodological worry is that auctorial commentaries may have a priming effect on the readers, for whom it can then be difficult to form independent opinions about the vignette stories. Motivated by this concern, a team of philosophers and psychologists has recently undertaken the task of checking whether it is really *people's* willingness to attribute knowledge that varies with stakes if "people" is meant to refer not only to (possibly carefully primed) professional philosophers but also to ordinary people (May et al. 2010). The team concluded that the aforementioned view is very probably false. That is certainly progress. If it takes the emergence of experimental philosophy to conduct this kind of research, then that will already make experimental philosophy a worthwhile new development.

[3] Callebaut (1993) gives an excellent overview of the naturalization program in philosophy up to the 1980s.

It is less clear to us that all that is currently undertaken under the banner of experimental philosophy targets the right question with the right methodology. The work just cited points toward a particular application of empirical research in philosophy, to wit, as offering a sort of check against unrestrained appeals to intuition. At the same time, there seems to be something like a neo-romanticist trend in experimental philosophy that we are more doubtful of. Just as the romantics thought that the enlightenment ideals, rather than leading to clearer insights, had alienated us from what was really true and valuable, and that the educated could learn about those matters again only from the common folks, some experimental philosophers seem to be driven by the idea that our intuitions have been corrupted by years of indoctrination in graduate school, and that we will have to tap into the reservoir of the folk's allegedly untainted intuitions about knowledge, or causation, or responsibility, or what have you, if we are to arrive at the true analysis of these concepts. Whatever romanticism may have brought us—and it certainly was not all good—we fear that neo-romanticist experimental philosophy will merely dredge up the kind of confusion about these concepts that graduate school training was meant to extirpate in the first place.

It is to be stressed that, as far as we are concerned, there is nothing intrinsically objectionable about the idea of philosophers engaging in science, provided they are willing to acquire the necessary additional skills or—better still, *and*—to look for experimental psychologists willing to work with them. Operationalizing research questions, developing proper task material, carrying out experiments under carefully controlled conditions, and properly analyzing and interpreting the data, are skills whose acquisition requires time and effort. It seems impossible to capture good experimental practice in a simple protocol that only has to be executed. Psychologists have discovered some common pitfalls in experimental design, but there are no general recipes for circumventing them. It is, in other words, not enough simply to know about the existence of such pitfalls, some of which have recently also been discussed in the philosophical literature (see, e.g., the methodological critique of philosophical survey-research in Cullen (2010) or the discussion of artifacts produced by uncontrolled use of abstract/concrete task material in Gendler 2007). One must also know how to deal with them, and part of this knowledge may be "tacit" in the sense of Polanyi (1976). To illustrate, we point to some pitfalls that are especially relevant to our program as well as to experimental philosophy.

The belief bias (Evans et al. 1994) is a well-known psychological effect that occurs if a participant ignores the premises and evaluates a conclusion (or the validity of an inference) only on the basis of the a priori believability of the conclusion. To avoid this pitfall, the experiment's instructions need to convey clearly to the participants that they are asked to evaluate the conclusion in light of the premises. If the task requires retrieval of background knowledge, the experimenter should control for it. This may be done by testing the experimental material with an independent sample before the experiment. But the finer details of this procedure may vary from one experiment to another, and learning how to carry out the procedure successfully requires prolonged and guided training.

To mention another potential problem, recent data suggest that experiments themselves may occasion certain "interpretation shifts" in participants. For example, it was found that, in a conditional interpretation task, a considerable number of participants started by giving a conjunction interpretation to the conditional but shifted their initial interpretation to the conditional event interpretation after solving the same type of task several times (Fugard et al. 2011a, b).[4] This finding suggests not only that experimenters would do well to carry out replication studies to investigate the stability of their results, but also would do well to control for within participants shifts in interpretations. Again, it is hard to provide general guidelines for working out the details of such checks; these details have to be determined on a case-by-case basis.

More generally, it cannot be emphasized enough that careful experimental conditions are a necessary ingredient of good methodological practice. For exploratory investigations almost anything may go, but good empirical research should strive for experiments that are in all respects well designed, from properly crafted stimuli to creating settings that allow participants to engage in the tasks and reasoning, from knowing when to use coders to knowing which statistical techniques to apply and how to apply them. It is our firm conviction that going to one's local McDonald's and register there willing customers' responses to Gettier cases, or Cohen's famous airport cases (Cohen 1988), or DeRose's equally famous bank cases (DeRose 1992), is decisively not the right approach—contrary to what some experimental philosophers seem to believe.

This is not to suggest that it is mainly the philosophers who stand to gain from collaboration with psychologists. We just mentioned the importance of presenting participants with properly crafted stimuli. Among other things, this means the stimuli should be free from ambiguities (unless one's research question requires such ambiguities, of course). Philosophers trained in the analytic tradition tend to be particularly good at detecting ambiguities and are generally sensitive to subtle differences in meanings and interpretations.[5]

Above all, however, philosophy is a real treasure trove of interesting empirical research questions. The previously mentioned area of conditionals is again a particularly good example. For decades now, philosophers have been working on the semantics and pragmatics of conditionals purely on the basis of a priori considerations, even though much that has come out of this work has testable consequences. Psychologists have only recently become aware of this, and have begun testing these

[4] Specifically, a bit more than half of the participants interpret the conditional as a conditional probability and a sizable proportion of participants interpret it as a conjunction in the beginning of these experiments. This replicates a robust finding in traditional probabilistic truth table tasks (Evans et al. 2003; Oberauer and Wilhelm 2003). However, in the end of the experiment, more than 80 % of the responses were consistent with the conditional probability interpretation. In these tasks inter-individual differences do not imply different reasoning/interpretation strategies but rather reflect how fast participants are able to adopt the competence interpretation. Although we doubt that the iterative presentation of the same task will resolve all aspects related to the deep problem of individual differences, we believe that it helps resolve at least some of them.

[5] An example of the latter kind is the distinction between reasoning *to* an interpretation and reasoning *from* an interpretation, where the former involves interpreting the meaning of the task material, while the latter begins only after the interpretation of the task material has been fixed by the participant and he or she starts to solve the task. We like to think that it is not coincidental that this distinction is due to recent joint work by a psychologist (or at least cognitive scientist) and a philosopher, to wit, Keith Stenning and Michiel van Lambalgen (Stenning and van Lambalgen 2008).

consequences, with sometimes quite surprising outcomes. Nonmonotonic reasoning is another case in point where philosophers and also computer scientists have developed inferential systems with clearly testable implications but have not cared to actually test these implications. This, too, is presently changing. In Sections 3–6, we want to discuss recent research on conditionals and also, more briefly, recent research on nonmonotonic reasoning as illustrative examples of work at the intersection of formal epistemology and experimental psychology, work that, we believe, can be regarded as signalling the emergence of the new program. Along the way, we show that the above recommendations for experimental philosophers are already implemented in this program. Prior to that, we describe another development in psychology and philosophy that has helped shape the research program we are presenting.

2 The Probabilist Turn

We have already noted that the new program is best thought of as a joint project of philosophers—in particular, formal epistemologists—and psychologists—in particular, experimental psychologists of reasoning. The advent of experimental philosophy, as described in the previous section, may have helped somewhat in making philosophers generally aware that experimental methods might be used to address philosophical problems as well. In greater part, however, we believe that the emergence of the new program is to be credited to the fact that—presumably purely coincidentally, and more or less at the same time—probability theory came to be broadly regarded as an important tool in the study of human reasoning, both among epistemologists and among psychologists. Once the two research communities had discovered that they were in agreement in a methodologically key respect, the road to fruitful collaboration was open.

It is not so easy to trace the origins of the probabilist turn in epistemology. Van Fraassen (1989, p. 151) characterizes the probabilist approach to epistemology as an "underground epistemology," the beginnings of which he locates in the seventeenth century, in the writings of Pascal and the Port Royal logicians. From an epistemological standpoint, van Fraassen's characterization seemed entirely appropriate even in 1989, notwithstanding the important work of Ramsey and de Finetti in the 1920s and 1930s, and of Carnap, Jeffrey, and a few others some decades later, and notwithstanding that probabilist methods had already attained some prominence in the philosophy of science (Earman 1992; Jeffrey 2004; Howson and Urbach 2006).

In the 1990s, however, the underground epistemology came to form the backbone of what was soon to be called "formal epistemology", a (then) new brand of epistemology in which probabilistic and, more generally, formal tools are used to address what are for the most part traditional questions from mainstream epistemology.[6] Formal epistemology turned almost overnight into a high profile research area, one that

[6]For more on what formal epistemology is and how it differs from mainstream epistemology as well as from the probabilist underground epistemology that van Fraassen refers to, see (Douven 2013c).

even started to dominate some debates in epistemology. The first debate of which this was true was no doubt the debate about the nature and epistemic role of coherence. There had been a lot of talk in epistemology about how coherence amongst beliefs may determine the status of those beliefs *qua* being justified, but much of the discussion had been in terms too vague to allow any of the crucial questions to be settled—such as, most notably, whether coherence is truth-conducive—or even to be addressed with any hope of success. This changed radically when formal epistemologists started to publish definitions of coherence, all in probabilistic terms, that did allow some of the crucial questions to be settled (see, e.g., Bovens and Hartmann 2003; Douven and Meijs 2007). Soon thereafter much the same probabilistic machinery came to be used more generally in tackling questions concerning justification and evidence. The end of this development is not yet in sight. What already seems clear, though, is that probability theory has firmly established itself as an indispensable analytical research tool in philosophy, next to logic (even if perhaps not quite as widely applicable as logic).

As noted, there has been a parallel probabilist turn in psychology. Since the first psychological experiments on reasoning, psychologists had broadly taken classical logic for granted as providing the right rationality norm for investigating reasoning (Lindworsky 1916; Störring 1908). Classical logic did not only guide the evaluation of the rationality of human inference but also inspired psychological theory building: the theory of mental models (Johnson-Laird 1983), for example, is inspired by mathematical model theory. Equally, the theory of mental rules (Rips 1994) and mental logic theory (Braine et al. 1998) are inspired by classical proof theory. Moreover, ideas from classical logic suggested many experimental paradigms like Wason's selection task and experiments on Aristotelian syllogisms and similar basic argument forms.

Uncertainty, however, used to be predominately investigated by the judgment and decision making (JDM) community. Kahneman and Tversky's *heuristics and biases* approach is one of the most prominent ones in this area (Kahneman et al. 1982). Whereas psychologists of reasoning focused on reasoning processes and logical validity judgments, proponents of the JDM community focused on judgments, decisions, and resulting outcomes. The main rationality norms for JDM research were probability theory and decision theory, but not classical logic. These differences in the focus of interest and in the rationality norms explain why there has been relatively little contact between the two communities.

But this may change, now that more and more psychologists of reasoning have started to adopt probabilistic approaches as well (Evans 2012; Evans and Over 2004; Oaksford and Chater 2007, 2009; Over 2009). The importance of conditional probability, for example, became apparent in the probabilistic truth table tasks (Evans et al. 2003; Fugard et al. 2011a; Oberauer and Wilhelm 2003), which in turn inspired new psychological models of conditionals—as will be amply illustrated in the following. This development has already led to closer contacts between psychologists of reasoning and formal epistemologists—resulting in the new program—but may in the future also lead to more collaborative work between psychologists of reasoning and researchers from the JDM camp.

3 Paradoxes of the Material Conditional

The material conditional $A \supset B$ is false if, and only if, the antecedent A is true and the consequent B is false; for all other possible truth value assignments, $A \supset B$ is true. The definition of the material conditional validates a number of inferences that appear counterintuitive under some natural language instantiations. For example, supposing the natural language conditional to have the semantics of $\supset$—as is claimed by the material conditional account. Then, B logically implies *If A, then B*. (We refer to this argument form as "Paradox 1".) The latter is also logically implied by $\neg A$. (We refer to this as "Paradox 2".) Intuitively, it seems counterintuitive to infer a conditional only on the basis of either its consequent or its negated antecedent. A fortiori, even if A and B are incompatible both argument forms are logically valid: "Carlos Slim Helú is a billionaire" logically implies "If Carlos Slim Helú is bankrupt, then Carlos Slim Helú is a billionaire" (Paradox 1). The latter conditional is also logically implied by "It is not the case that Carlos Slim Helú is bankrupt" (Paradox 2).

These counterintuitive consequences of the material conditional account led to a variety of responses.[7] For instance, some have appealed to Gricean maxims of conversation (Grice 1975) to explain away the paradoxes along pragmatic lines, pointing out that an assertion of "If A, then B" only on the basis of one's knowledge of either B or $\neg A$ would violate the Gricean cooperative principle, which requires, among other things, that one communicate in a brief and informative way.

Gricean approaches were also adopted by psychologists. Most prominently, the theory of mental models claims that the paradoxes "throw semantic information away", as "the premises rule out more possibilities than do their conclusions" (Johnson-Laird and Byrne 2002, p. 652). The premise of Paradox 1 (B), for example, rules out the possibility $\neg A \wedge \neg B$, which is (uninformatively) introduced by the conclusion ($A \supset B$). The theory predicts that Paradox 1 is endorsed, if the possibility $\neg A \wedge \neg B$ is ruled out by world knowledge and thus cannot be (uninformatively) introduced by the conclusion. This account was criticized by Bonnefon and Politzer, who impose two (necessary and sufficient) pragmatic conditions on the acceptability of Paradox 1 (Bonnefon and Politzer 2011, p. 146):

BP1: The truth of the antecedent A has bearings on the relevance of asserting the consequent B.

BP2: The speaker can reasonably be expected not to be in a position to assume that A is false.

Mental probability logic provides an explanation of the paradoxes in purely probabilistic terms without involving pragmatics (Pfeifer 2013a, b; Pfeifer and Kleiter 2011). In this approach, the premise set may contain probabilistic and/or logical information. The inference consists in transmitting the probability of the premise set deductively to the conclusion. Moreover, conditionals are interpreted as conditional events. A conditional event $B \mid A$ is true if A is true and B is true, false if A is true and

[7] For an overview, see Bennett (2003).

B is false, and *undetermined* if A is false. Paradox 1 is then interpreted as an inference from a marginal probability to a conditional probability: For all[8] probability values x:

$$\text{From } \Pr(B) = x \text{ infer } 0 \leqslant \Pr(B \mid A) \leqslant 1 \text{ is coherent.}$$

It is easy to see that the paradox is blocked here: from $\Pr(B)$ one cannot infer anything about $\Pr(B \mid A)$. The same is true for Paradox 2: from $\Pr(\neg A)$ one cannot infer anything about $\Pr(B \mid A)$. If the conditional in the conclusion is interpreted as a material conditional, then both argument forms are probabilistically informative: for both paradoxes, the tightest coherent probability bounds of $A \supset B$ are the premise probability (lower probability bound) and 1 (upper probability bound).

A recent empirical study of both paradoxes shows that the clear majority of participants responds by probabilistically non-informative response patterns (Pfeifer and Kleiter 2011).[9] This corresponds to the conditional event interpretation of indicative conditionals, and speaks against the hypothesis that people interpret indicative conditionals as material ones. The task material was formulated in a neutral way without suggesting probabilistic or logical (in) dependencies between the involved terms. Here, both paradoxes are blocked because they are probabilistically non-informative.[10]

What happens if we know something about the logical and/or probabilistic relations between A and B? In the above case of Carlos Slim Helú, for instance, we know that being bankrupt is incompatible with being a billionaire. Thus, adding this information to the premise set yields the following intuitively plausible inference, where the conclusion obtains a low probability value:

$$\text{From } \Pr(B) = 1 \text{ and } A \wedge B \equiv \bot \text{ infer } \Pr(B \mid A) = 0 \text{ is coherent.}$$

Here, the zero probability of the conclusion shows that the the paradoxical inference is blocked. Likewise, the intuitive plausibility of the inference

> From "Carlos Slim Helú is a billionaire" infer "If Carlos Slim wins a billion in the lottery, then Carlos Slim Helú is a billionaire".

[8] "All" here includes the special case where $\Pr(B) = 1$. Bonnefon and Politzer correctly point out that in this case " 'If x, y' must also be certain, and the inference is valid" (Bonnefon and Politzer 2011, p. 154). This is true for the standard approach to probability, which defines the conditional probability $\Pr(B \mid A)$ by the fraction of the joint and the marginal probability, $\Pr(A \wedge B)/\Pr(A)$ (provided $\Pr(A) > 0$, otherwise $\Pr(B \mid A)$ is undefined). In the framework of coherence-based probability theory, however, $\Pr(B \mid A)$ is not necessarily certain if $\Pr(B) = 1$. As $\Pr(A)$ may be equal to 0, it follows that $\Pr(B \mid A)$ may also be equal to 0, and therefore $0 \leqslant \Pr(B \mid A) \leqslant 1$ is coherent (Pfeifer 2013a).

[9] An anonymous referee noted that it is not clear whether the observed response—that nothing follows—fully accounts for a "subjective feeling of oddity", which emerges from some instantiations of the paradoxes of the material conditional. The tasks used in Pfeifer and Kleiter (2011) were formulated by neutral instantiations of A and B (i.e., in terms of colors and figures), which means that a feeling of oddity—if it occurred—was based on the formal structure of the paradoxes only.

[10] An argument form is *probabilistically non-informative* iff the tightest coherent probability bounds on the conclusion are zero and one, respectively, under all possible probability values of the premises (Pfeifer and Kleiter 2009).

can be rationally reconstructed by adding the logical constraint that winning a billion in the lottery logically implies being a billionaire:

$$\text{From } \Pr(B) = 1 \text{ and } A \supset B \equiv \top \text{ infer } \Pr(B \mid A) = 1 \text{ is coherent.}$$

Note that in both reconstructions, the logical constraint is sufficient to obtain the conclusion's probability.

Another and weaker way to obtain a probabilistically informative version of Paradox 1 is to use the intuition of BP2, that is, the reasoner is not in a position to assume that the antecedent A is false. This can be interpreted in probabilistic terms such that $\Pr(A)$ is greater than .5. Thus, assuming "high" premise probabilities, $\Pr(A) > .5$ and $\Pr(B) > .5$, we obtain the following probabilistically informative inference:[11]

$$\text{From } \Pr(B) = x \text{ and } \Pr(A) = y \text{ infer}$$

$$\frac{x+y-1}{y} \leqslant \Pr(B \mid A) \leqslant \min\left\{\frac{x}{y}, 1\right\} \text{ is coherent.}$$

Probabilistic non-informativeness is sufficient to block Paradox 1 and 2, but also to block Premise strengthening and Contraposition: empirical evidence suggests that people understand that one cannot infer probabilistically informative interval in these argument forms (Pfeifer and Kleiter 2010).

Not all inferences can be modelled in purely probabilistic terms. The hypothetical syllogism,[12] for example, is probabilistically non-informative but most people draw probabilistically informative inferences, which are consistent with the following argument form (Pfeifer and Kleiter 2010):[13]

$$\text{From } \Pr(B|A) = x \text{ and } \Pr(C|A \wedge B) = y \text{ infer}$$

$$xy \leqslant \Pr(C \mid A) \leqslant xy + 1 - x \text{ is coherent.}$$

Mental probability logic explains this by conversational implicatures (Pfeifer and Kleiter 2010): People add the antecedent of the first premise of the hypothetical syllogism to the antecedent of the second premise.

Enriching the premise set of the paradoxes by logical and/or probabilistic constraints generates fruitful new empirical hypotheses. This illustrates that the formal analysis provides clear psychological predictions. Moreover, the results of empirical work can feed back into the formal analysis and into rational reconstruction of the data; thereby probabilistic interpretations of the paradoxes of the material conditional provide a good illustration of the new program.

---

[11] We observe that this argument form is a special case of the cautious monotonicity rule of System P. (Compare the probability propagation rules in Gilio (2002)).

[12] From "If A, then B" and "If B, then C" infer "If A, then C".

[13] This argument form corresponds to the cut rule of System P (Gilio 2002). Moreover, it can be interpreted as a conditional version of modus ponens: each premise and the conclusion conditionalize on A, if A is dropped, the modus ponens remains.

4 Nonmonotonic Reasoning

The same holds true of recent work on nonmonotonic reasoning; this, too, illustrates how empirical work can inform formal work and how formal work can inspire empirical research.

Nonmonotonic reasoning originated as a subfield of artificial intelligence concerned with the study of formal structures governing the rational retraction of conclusions in the light of new evidence. The paradigm example is the Tweety case. From the two premises “Tweety is a bird” and “Birds can fly” it seems reasonable to infer that Tweety can fly. If the premise set is augmented by the further premise “Tweety is a penguin”, common sense tells us to retract the conclusion that Tweety can fly, as penguins cannot fly. In the framework of classical logic, however, augmenting the premise set of a logically valid argument cannot lead to a logically invalid argument: classical logic is *monotonic*. Nonmonotonic formalisms were developed to account for situations like the Tweety case.

Nonmonotonic reasoning can be conceived as a promising rationality framework for studying human inference. Besides, much of the formal work is motivated by the observation that real agents do reason nonmonotonically; see, for instance, (McCarthy 1977, p. 1040) and (Brewka 1991, pp. 2, 13). Pelletier and Elio (1997) even go so far as to argue that nonmonotonic reasoning is a genuinely psychologistic endeavor:

> [C]onsidering how people actually do default reasoning is an important and necessary grounding for the entire enterprise of formalizing default reasoning. (Pelletier and Elio 1997, p. 165)
>
> We have claimed in this paper that, unlike classical logic, default reasoning is basically a psychologistic enterprise. (Pelletier and Elio 1997, p. 177).

We do not fully agree with this assessment, however. Specifically, it seems to us that, given that purely a priori considerations gave rise to a vast literature on nonmonotonic formalisms (see, e.g., Antoniou 1997; Brewka et al. 2011), nonmonotonic reasoning should not be characterized as a psychologistic endeavor.

Be this as it may, while many systems characterize nonmonotonicity as absence of the monotonicity property, some systems attempt to describe nonmonotonic reasoning positively. One of the most important ones is System P (Kraus et al. 1990). This system describes a set of basic rationality postulates that any system of nonmonotonic reasoning should satisfy. System P not only guarantees nonmonotonicity but also describes nonmonotonic reasoning positively in an axiomatic way. The key term of this system is the nonmonotonic consequence relation, denoted by $\mid\!\sim$, where $A \mid\!\sim B$ is standardly interpreted as meaning that if A, then *normally* B. Various semantics have been proposed for System P. Among them, at least two deal with uncertainty: possibility semantics (Benferhat et al. 1997) and probability semantics (e.g., Adams 1975; Gilio 2002). Rather than being based on empirical work, both semantics inspired such work.

Possibilistic semantics flesh out the interpretation of $A \mid\!\sim B$ in terms of the possibility of $A \wedge B$ being greater than the possibility of $A \wedge \neg B$. This semantics received empirical support from work reported in Benferhat et al. (2005) and Da Silva Neves et al. (2002). The probability semantics, which is of special interest in the present context, interprets $A \mid\!\sim B$ in terms of high conditional probability $\Pr(B \mid A)$ (e.g., Gilio 2002; Hawthorne and Makinson 2007; Schurz 2005). In Gilio's coherence-based probability semantics, the rules of System P are conceived as argument forms where the premises are assigned (imprecise) probabilities and the uncertainty is transmitted deductively to the conclusion. The tightest coherent probability bounds on the conclusion are obtained by probability propagation rules.

Gilio's coherence-based probability semantics of System P was investigated in a series of psychological experiments (Pfeifer and Kleiter 2003, 2005, 2006). People tend to endorse the nonmonotonic rules of System P, but not the monotonic counterparts *monotonicity*—from $\Pr(E_3 \mid E_1) = x$ infer $\Pr(E_3 \mid E_1 \wedge E_2) \in [0, 1]$—and *contraposition*—from $\Pr(E_2 \mid E_1) = x$ infer $\Pr(\neg E_1 \mid \neg E_2) \in [0, 1]$. While these monotonic counterparts are probabilistically non-informative, all System P rules are probabilistically informative. This situation is comparable to the paradoxes of the material conditional discussed in Section 3: The probabilistic non-informativeness blocks the paradoxes and the monotonicity property, which is consistent with philosophical intuitions and human thought. The mentioned series of experiments on System P offers a clear example of formal work that stimulated empirical research on nonmonotonic reasoning. Moreover, the experimental paradigm developed for these experiments showed for the first time how people deal with probability intervals and also showed that most people do not violate the probability axioms under careful experimental conditions.

Even if nonmonotonic reasoning should not be characterized as a psychologistic endeavor, it *is* true that empirical research on nonmonotonic reasoning also inspired formal work. Ford's System LS (Ford 2004), for instance, is inspired by her empirical work on nonmonotonic reasoning strategies. System LS deals with three levels of strength between the antecedents and consequents of nonmonotonic consequence relations, which are rationally justified by Gilio's probability semantics of System P (Gilio 2002). Here, the strongest one, $A \mid\!\sim_1 B$, means that more than half of the A are B; the intermediate one in strength, $A \mid\!\sim_2 B$, means that some A are B; and the weakest one, $A \mid\!\sim_3 B$, means that a relationship between A and B has been derived but it is possible that no A are B (Ford 2004, 106). The rules of System LS are qualitative and not explicitly probabilistic. System LS allows for rationally reconstructing some human reasoning strategies which cannot be described within a purely monotonic framework. While doing justice to human reasoning strategies, System LS solves nonmonotonic benchmark problems, like the Tweety case, at the formal level as well.

Finally, there are examples of psychological experiments which investigate nonmonotonic reasoning without explicit reference to a normative theory. The suppression effect (Byrne 1989) is a case in point. It demonstrates that the endorsement rate of valid inferences (such as modus ponens) can decrease substantially if an additional premise is added to the premise set. However, without the reference to formal work, this effect is no more than an observation and—even worse—may misinterpret human reasoning strategies as being irrational, in view of the fact that the monotonicity of classical logic is violated. It took more than a decade for the suppression effect to be investigated within probabilistic (Politzer and Bourmaud 2002; Politzer 2005) or nonmonotonic (Stenning and van Lambalgen 2005) frameworks, which allowed for rational justifications of the data.

5 The Assertability / Acceptability of Conditionals

Not so long ago, it was widely believed that assertability goes by probability, meaning that how assertable one deems a given sentence is measured by the probability one assigns to that sentence. This was meant to hold generally, for non-conditional and conditional sentences alike. While the proposal seems simple enough, various philosophers realized that it raises a question that is far from simple to answer, to wit, the question of what the probability of a conditional is. Probability theory tells us how to calculate the probability of, say, $\neg A$ on the basis of the probability of A, or the probability of $A \wedge B$ on the basis of the probability of A and the probability of B given A. But it is unclear whether probability theory says anything at all about the probability of the conditional "If A, B". The reason for the lack of clarity is, of course, the lack of a generally agreed-upon semantics of conditionals. If, as the material conditional account holds, conditionals have the truth conditions of the corresponding material conditionals, then the probability of "If A, B" is obviously, and quite simply, equal to $\Pr(\neg A \vee B) = 1 - \Pr(A \wedge \neg B)$. But this would have the consequence that all conditionals with improbable antecedents are probable, and it is not hard to come up with natural language conditionals with improbable antecedents that are pre-theoretically highly improbable. There are other reasons to be wary of the material conditional account, as we saw in Section 3.

A more plausible answer, prima facie, to the question of the probability of conditionals is that $\Pr(\text{If } A, B) = \Pr(B \mid A)$. Indeed, in the past decade experimental psychologists have tested this answer several times over, and all experiments so far have confirmed its empirical adequacy (e.g., Over and Evans 2003; Evans and Over 2004; Fugard et al. 2011a; Over et al. 2013). But rather than settling the question, these results raise a puzzle. For it had been argued by Lewis and others that the probability of a conditional simply cannot equal the corresponding conditional probability, on pain of triviality. Specifically, a number of formal results—so-called triviality arguments—seemed to show that the equality can only hold for probability functions that have various features which make them unsuited for representing the degrees of belief of real human agents. As the hedged wording ("seemed to show") suggests, we believe that these triviality arguments may not be as inescapable as they are almost universally believed to be (Douven 2013a; Douven and Verbrugge 2013). Nevertheless, here we sidestep possible worries one may have about these arguments and focus on a popular response that takes the arguments at face value.

What many now regard as the right response to the triviality arguments really predated them. Skeptical about the idea that conditionals express propositions, Adams (1965)—more than ten years prior to the publication of the first triviality results (Lewis 1976)—had already proposed that conditionals do not have probabilities of truth *strictu sensu*, but that they do have degrees of assertability (or degrees of acceptability, the term he later came to prefer), and that those are measured by the corresponding conditional probabilities. More exactly, the proposal was this:

$$\textbf{Adams' Thesis (AT)}\ \mathrm{Acc}(\text{If } A, B) = \Pr(B \mid A), \text{ provided } \Pr(A) \neq 0;$$
$$\text{if } \Pr(A) = 0, \text{ then } \mathrm{Acc}(\text{If } A, B) = 1.$$

Here, "Acc(·)" denotes the degree of assertability / acceptability of a sentence. It is also to be noted that (AT) pertains only to so-called simple conditionals, that is, conditionals whose antecedents and consequents are not themselves conditional in form nor consist of compounds at least one of whose component sentences is a conditional. This restriction to simple conditionals is inescapable if conditionals do not express propositions. For let A or B be a conditional. Then, given the standard ratio definition of conditional probability, the probability of "If A, B" would equal $\Pr(A \wedge B)/\Pr(A)$, according to (AT). But if conditionals do not express propositions, the numerator of this ratio is not well-defined, for then conditionals cannot occur in conjunctions; the conjunction operator is a *propositional* operator, after all.[14]

While some have complained that (AT) is exceedingly weak, precisely because of the aforementioned restriction (McGee 1989), no one has doubted its empirical adequacy. Indeed, the literature is awash in statements expressing the obvious correctness of (AT). Perhaps this is not so surprising: the thesis does have an intuitive ring to it. On the other hand, philosophers' intuitions have been proven wrong more than once—an observation that we identified earlier as a main motivation of the experimental philosophy movement—and so, given that (AT) has been around for so long, and given also that it plays a pivotal role in the theorizing about conditionals, one would have expected to find some experimental work on it in the literature. Surprisingly, the first experiments addressing (AT) were not carried out until very recently, and have been reported in Douven and Verbrugge (2010).

Because Adams' thesis is sometimes said to express the idea that the assertability/ acceptability of a conditional *goes by* the corresponding conditional probability, Douven and Verbrugge in effect investigate, along with the empirical adequacy of (AT), the empirical adequacy of a family of weaker versions of that thesis that could all be said to explicate the designated idea, ranging from the strict thesis (AT) to the much weaker claim that the degrees of acceptability of conditionals and the corresponding conditional probabilities are at least moderately correlated.

Douven and Verbrugge's stimuli consisted of what linguists (e.g., Dancygier 1998; Declerck and Reed 2001) have called "inferential conditionals", that is, conditionals reflecting an inference from the conditional's antecedent to its consequent. Various typologies of such conditionals already exist in the linguistics literature, but Douven and Verbrugge propose a new typology, based on the type of inference a conditional reflects—whether deductive, inductive, or abductive—where the type may depend on the background assumptions being made in the context in which a conditional is asserted or is being evaluated.

[14] If conditionals do not express propositions, they cannot occur in *ordinary* conjunctions. They *can* occur in what Adams calls "quasi-conjunctions" (Adams 1975, p. 46 f). By Adams' own admission, however, quasi-conjunctions lack some important logical features of conjunctions.

In their main experiment, Douven and Verbrugge asked one group of participants to judge the degree of acceptability of 30 inferential conditionals—10 deductive inferential (DI) conditionals, 10 inductive inferential (II) conditionals, and 10 abductive inferential (AI) conditionals—and another group to judge the corresponding conditional probabilities. Douven and Verbrugge obtained mixed, but mostly negative, results. Taken over all inferential conditionals, the participants' responses manifestly refuted (AT) and also various weaker theses; the strongest thesis that was supported was that acceptability of conditional and conditional probability are highly correlated. However, when the types of inferential conditionals were considered separately, the results were partly better, but partly even worse. They were better insofar as (AT) was found to hold for DI conditionals. For AI conditionals, the result was the same as for inferential conditionals generally, in that the high correlation thesis was the strongest thesis found to hold for AI conditionals. However, for II conditionals, the results were worse: there was at best a moderately high correlation between acceptability of conditional and conditional probability.[15]

We take this research to constitute a clear point in favor of the new program. Many may be somewhat disappointed that the story to be told about the assertability / acceptability of conditionals is not quite as simple as Adams thought, as did many others who followed him. But at least the story we now have is based on empirical evidence rather than on casual introspection and philosophical speculation.

6 Updating on Conditionals

Probably, the greatest part of the written or spoken information that we receive comes to us in the form of factive sentences; we read or are told that there will be a football match tomorrow afternoon, or that John has an important exam next week. If the information is new to us, we learn something, and we may, on that basis, come to believe certain things that we previously disbelieved and we may also revise some beliefs we hold. But some information that we receive comes in an explicitly conditional form. We may learn, for instance, that if it continues to rain, then tomorrow's match will be cancelled, or that if John fails next week's exam, he will stay down a grade. If any of these things are new to us, we would also seem to learn something,[16] and we might reasonably come to believe certain things or revise previously held beliefs in light of the newly acquired information. For example, upon learning that if John fails next week's exam, he will stay down a grade, one might revise one's belief

[15] Note that, in view of the experiments on the probabilities of conditionals mentioned in the text, Douven and Verbrugge's results also show that the probability of a conditional is not the same as the degree of acceptability of a conditional, *pace* Adams.

[16] An anonymous referee noticed that the fact that we seem to learn something in such cases puts pressure on Adams' non-propositional view of conditionals. We agree, albeit only insofar as Adams' view rules out the seemingly most straightforward type of proposal for modelling conditional updates, to wit, those proposals that model such updates as the accommodation, in some way, of the proposition (putatively) expressed by the conditional. But there are other forms that updating on a conditional might take; see Douven and Romeijn (2011) and Douven (2012).

that John had done well on his other exams this year (perhaps reasoning that if he really had, that would compensate for a possible failure next week).

Most people who have thought about the issue of belief change would agree that by now we have a fairly firm grip on how we adapt our beliefs on the basis of factive information that we learn. Work done in the tradition of what is commonly known as AGM theory (cf. Gärdenfors 1988 for an overview) has helped in this respect, but we submit that, especially, developments in the Bayesian tradition, both by philosophers and by psychologists, have led to important progress toward understanding formally and empirically the finer mechanics of belief change. Things look very differently when it comes to understanding how we adjust our beliefs in response to the receipt of conditional information. While the epistemic processing of conditional information seems a routine matter, and while we also have clear intuitions about how learning particular pieces of conditional information should affect a person's belief system, there exist few (if any) working models for such learning events. There is some AGM-style work on updating on conditionals (e.g., Boutilier and Goldszmidt 1995; Kern-Isberner 1999), but issues of empirical adequacy of the proposals that have come out of this—do they make the right predictions about how people actually respond to the receipt of conditional information?—have still to be considered. Also, this work is not aimed at answering the question of how we update our degrees of belief on conditional information. What is more surprising is that neither will one find an answer to that question in the Bayesian literature, where, in fact, that question has hardly ever been raised.

The reason for this cannot be that, from a Bayesian perspective, updating on conditionals is business as usual. According to Bayesians, rational people update on incoming information by conditionalizing on it. But it is unclear what it is to conditionalize on a conditional, and this has long been known (see Skyrms (1980), p. 169). There has been some recent progress on this front: in Douven and Romeijn (2011), a procedure is described for updating on conditional information that satisfies several intuitive desiderata for such updates. However, Douven and Romeijn's proposal has rather limited scope in that it pertains only to the special type of case in which learning of the conditional ought not affect one's probability for the antecedent of the conditional. Moreover, while updating probabilistically on conditionals is clearly a process governed by some norms—for instance, we have clear intuitions about whether, given a context, an update on a particular conditional should lead to an increase or rather a decrease in the probability we assign to the conditional's antecedent, or whether it should leave that probability unaffected—Douven and Romeijn do not provide a systematic account of what determines which of the three is the rational thing to do. Indeed, they explicitly refrain even from offering a systematic account of the circumstances in which it would be reasonable to opt for the special learning rule they propose.

The model of updating on conditional information presented in Douven (2012) is more specific in this regard. Underlying that model is the thought that, at least for a large class of conditionals, learning of a conditional is accommodated on the basis of explanatory considerations. To be more exact, the proposal is that if learning a given conditional has no effect on the explanatory status of the antecedent of that conditional—it does not make the antecedent a better or worse explanation for something one knows or at least believes to be the case—then it should not alter the antecedent's probability; in that case, Douven and Romeijn's rule for this type of case applies. If, on the other hand, the learning event does have an effect on the explanatory status of the antecedent, then the probability for the antecedent should change: upwards if the antecedent's explanatory status has gone up due to the learning of the conditional, and downwards if its explanatory status has gone down.[17]

The proposal is in line—and is meant to be in line—with the intuition, voiced by many, that explanatory considerations play a fundamental role in probabilistic reasoning (see, e.g., Lipton 2004; Weisberg 2009). But however plausible this thought may appear, the exact role of explanatory considerations in probabilistic reasoning still remains to be clarified (Douven 2011, 2013b). As Douven (2012) argues, if the explanatory model of updating on conditionals can be shown to be correct, then at least one specific role—perhaps the main role—of explanatory considerations will have been pinpointed.

However, the explanatory model for updating on conditionals is still severely limited. Except for the type of case in which learning a conditional does not change the explanatory status of the conditional's antecedent, the model does not give, for every proposition, the new probability one ought to assign to it after learning a conditional—contrary to what one would wish. In fact, for the cases in which learning a conditional changes the explanatory status of the antecedent, the model only issues qualitative prescriptions, and then only for how to change one's probability for the antecedent of the conditional one has learned, not for other propositions.

This is unfortunate. For instance, with no special assumptions about the context of utterance, an utterance of "If it starts raining, the match will be cancelled" by a person one trusts on the matter is likely to raise one's probability that if it starts snowing, the match will be cancelled; in any event, there would seem nothing rationally untoward in such a response to the designated utterance. But the explanatory model is silent on such belief changes, and it seems that presently no one has offered a way to account for them.

In previous sections, we described work that has recently been done or is ongoing and which falls squarely in the new program. To the best of our knowledge, no one, whether in philosophy or in psychology, is currently working on updating on conditionals. We see the aforementioned problem—that no one seems to have an idea of what an even moderately general rule of updating on conditionals might look like—as a patent opportunity for collaborative research by experimental psychologists and formal epistemologists, and indeed as an opportunity to further support the case for the new program.

A first step toward realizing this goal would be to test the empirical adequacy of the explanatory model: even though, as indicated, it is predictively not as strong as one would like, it does make clear predictions which, moreover, are testable. For instance, Douven (2012) suggests a between subjects experiment which asks one

[17] For an important different approach to modelling updates on conditionals, see Hartmann and Rafiee Rad (2012).

group of participants to rate the explanatory status of a proposition *A* in the light of background knowledge provided by a vignette; asks a second group to rate that same status in an extension of the vignette in which a conditional "If *A*, *B*" has been learned, for some well chosen proposition *B*; asks a third group to rate the probability of *A* in the original vignette; and asks a fourth group to rate the probability of *A* in the extended vignette story. By repeating this for sufficiently many and sufficiently varied vignette–conditional pairs, and by verifying whether a correlation exists between the two variables, one may obtain evidence for the explanatory model. To describe the experiment in broad strokes is one thing; to determine a detailed design, and particularly to figure out the best way to elicit judgments about the explanatory statuses of propositions, is another, and requires specialized knowledge. Here, the help of experimental psychologists, preferably ones who have worked on the role of explanatory considerations in human reasoning (e.g., Koehler 1991; Lombrozo 2007), would seem indispensable.

A second and potentially more important step would involve looking for finer patterns in the data about updating on conditionals as obtained in the previously described experiment and ideally in other experiments as well. This might allow for some sort of reverse engineering. By getting some understanding of how people *actually* update on conditionals (or perhaps how cognitively superior people update on conditionals), one might come to see, at least in outline form, the norms that philosophers assume to govern such updates but to this date have been unable to pinpoint. Naturally, there is no guarantee that the data will reveal any stable patterns in how people adjust their probabilities in light of conditional information that they receive. But the only way to find out is to give it a try. If the attempt succeeds, that would be a striking illustration of how empirical data may stimulate and aid theory building in formal epistemology.

We have argued that a convergence in the interests and methodologies of formal epistemologists and psychologists of reasoning has given rise to a new field of research. The case studies we presented were intended to illustrate the point that, by joining forces, these two communities have achieved understanding of uncertain reasoning—including reasoning about and involving conditionals—that each community probably could not have achieved on its own. Thus, an interdisciplinary position is of vital importance for future work. We have also pointed to some open questions that clearly call for a similar combined approach. This justifies the hope that the research program we outlined is going to be a lively research paradigm for many years to come.

Acknowledgments The authors thank two anonymous reviewers and Paul Égré for useful comments. This work is supported by the FWF project P20209 and the DFG grant PF 740/2-1 (project leader: Niki Pfeifer; project within the DFG Priority Program SPP 1516 "New Frameworks of Rationality"). Niki Pfeifer is supported by the Alexander von Humboldt Foundation.

References

Adams, E.W. 1965. The logic of conditionals. *Inquiry* 8: 166–197.
Adams, E.W. 1975. *The logic of conditionals*. Dordrecht: Reidel.
Antoniou, G. 1997. *Nonmonotonic reasoning*. Cambridge: MIT Press.

Benferhat, S., J.-F. Bonnefon, and R. Da Silva Neves. 2005. An overview of possibilistic handling of default reasoning, with experimental studies. *Synthese* 146: 53–70.

Benferhat, S., D. Dubois, and H. Prade. 1997. Nonmonotonic reasoning, conditional objects and possibility theory. *Artificial Intelligence* 92: 259–276.

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

Bonnefon, J.-F., and G. Politzer. 2011. Pragmatics, mental models and one paradox of the material conditional. *Mind & Language* 26: 141–155.

Boutilier, C., and M. Goldszmidt 1995. On the revision of conditional belief sets. In *Conditionals. From philosophy to computer science*, eds. G. Crocco, L. Fariñas del Cerro, A. Herzig, 267–300, Oxford: Oxford University Press.

Bovens, L., and S. Hartmann. 2003. *Bayesian epistemology*. Oxford: Oxford University Press.

Braine, M.D.S., and D.P. O'Brien. eds. 1998. *Mental logic*. Mahwah: Erlbaum.

Brewka, G. 1991. *Nonmonotonic reasoning: Logical foundations of commonsense*. Cambridge: Cambridge University Press.

Brewka, G., V.W. Marek, and M. Truszczyéski. 2011. *Nonmonotonic reasoning. Essays celebrating its 30th anniversary of studies in logic, vol. 31.* London: College Publications.

Byrne, R.M.J. 1989. Suppressing valid inferences with conditionals. *Cognition* 31: 61–83.

Callebaut, W. 1993. *Taking the naturalistic turn*. Chicago: University of Chicago Press.

Chater, N., and M. Oaksford. 1999. The probability heuristics model of syllogistic reasoning. *Cognitive Psychology* 38: 191–258.

Cohen, S. 1988. How to be a fallibilist. *Philosophical Perspectives* 2: 91–123.

Cullen, S. 2010. Survey-driven romanticism. *Review of Philosophy and Psychology* 1: 275–296.

Da Silva Neves, R., J.-F. Bonnefon, and E. Raufaste. 2002. An empirical test of patterns for nonmonotonic inference. *Annals of Mathematics and Artificial Intelligence* 34: 107–130.

Dancygier, B. ed. 1998. *Conditionals and predictions: Time, knowledge and causation in conditional constructions*. Cambridge: Cambridge University Press.

Declerck, R., and S. Reed. 2001. *Conditionals: A comprehensive empirical analysis*. Berlin: Mouton de Gruyter.

DeRose, K. 1992. Contextualism and knowledge attributions. *Philosophy and Phenomenological Research* 52: 913–929.

Douven, I. 2008. The evidential support theory of conditionals. *Synthese* 164: 19–44.

Douven, I. 2011. Abduction. In *The Stanford encyclopedia of philosophy,* ed. E.N. Zalta. Spring.

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

Douven, I. 2013a. The epistemology of conditionals. *Oxford Studies in Epistemology* 4: 3–33.

Douven, I. 2013b. Inference to the best explanation, Dutch books, and inaccuracy minimisation. *Philosophical Quarterly* 63: 428–444.

Douven, I. 2013c. Formal epistemology. In *Oxford handbooks online in philosophy*, ed. S. Goldberg. Oxford: Oxford University Press. Forthcoming.

Douven, I., and R. Dietz. 2011. A puzzle about Stalnaker's hypothesis. *Topoi* 30: 31–37.

Douven, I., and W. Meijs. 2007. Measuring coherence. *Synthese* 156: 405–425.

Douven, I., and J.-W. Romeijn. 2011. A new resolution of the Judy Benjamin problem. *Mind* 120: 637–670.

Douven, I., and S. Verbrugge. 2010. The Adams family. *Cognition* 11: 302–318.

Douven, I., and S. Verbrugge. 2013. The probabilities of conditionals revisited. *Cognitive Science* 37: 711–730.

Earman, J. 1992. *Bayes or bust? A critical examination of Bayesian confirmation theory*. Cambridge: MIT Press.

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

Evans, J. St B.T. 2012. Questions and challenges to the new psychology of reasoning. *Thinking & Reasoning* 18: 5–31.

Evans, J. St B.T., J.L. Allen, S. Newstead, and P. Pollard. 1994. Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology* 6: 263–285.

Evans, J. St B.T., S.J. Handley, and D.E. Over. 2003. Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29: 321–355.

Evans, J. St B.T., and D.E. Over. 2004. *If*. Oxford: Oxford University Press.

Ford, M. 2004. System LS: A three-tiered nonmonotonic reasoning system. *Computational Intelligence* 20: 89–108.

Fugard, A.J.B., N. Pfeifer, and B. Mayerhofer. 2011a. Probabilistic theories of reasoning need pragmatics too: Modulating relevance in uncertain conditionals. *Journal of Pragmatics* 43: 2034–2042.
Fugard, A.J.B., N. Pfeifer, B. Mayerhofer, and G.D. Kleiter. 2011b. How people interpret conditionals: Shifts towards the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37: 635–648.
Gärdenfors, P. 1988. *Knowledge in flux*. Cambridge: MIT Press.
Gendler, T.S. 2007. Philosophical thought experiments, intuitions, and cognitive equilibrium. *Midwest Studies in Philosophy* 31: 68–89.
Gilio, A. 2002. Probabilistic reasoning under coherence in System P. *Annals of Mathematics and Artificial Intelligence* 34: 5–34.
Grice, H.P. 1975 Logic and conversation. In *Syntax and semantics, Speech acts,* vol. 3, eds. P. Cole, J.L. Morgan. New York: Seminar Press.
Hahn, U., and M. Oaksford. 2006. A Bayesian approach to informal argument fallacies. *Synthese* 152: 207–236.
Harris, A.J.L., and U. Hahn. 2009. Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35: 1366–1373.
Hartmann, S., and S. Rafiee Rad. 2012. *Updating on conditionals = Kullback–Leibler + causal structure*. Manuscript.
Hawthorne, J., and D. Makinson. 2007. The quantitative/qualitative watershed for rules of uncertain inference. *Studia Logica* 86: 247–297.
Howson, C., and P. Urbach. 2006. *Scientific reasoning: The Bayesian approach*, 3rd edn. Chicago: Open Court.
Jackson, F. ed. 1991. *Conditionals*. Oxford: Oxford University Press.
Jeffrey, R. 2004. *Subjective probability: The real thing*. Cambridge: Cambridge University Press.
Johnson-Laird, P.N. 1983. *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.
Johnson-Laird, P.N., and R.M.J. Byrne. 2002. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review* 109: 646–678.
Kahneman, D., P. Slovic, and A. Tversky. eds. 1982. *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
Kauppinen, A. 2007. The rise and fall of experimental philosophy. *Philosophical Explorations* 10: 95–118.
Kemp, C., N.D. Goodman, and J.B. Tenenbaum. 2010. Learning to learn causal models. *Cognitive Science* 34: 1185–1243.
Kern-Isberner, G. 1999. Postulates for conditional belief revision. In *Proceedings of the 16th international joint conference on artificial intelligence*, ed. T. Dean, 186–191. San Francisco: Morgan Kaufmann.
Koehler, D.J. 1991. Explanation, imagination, and confidence in judgment. *Psychological Bulletin* 110: 499–519.
Kraus, S., D. Lehmann, and M. Magidor. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44: 167–207.
Lewis, D. 1976. Probabilities of conditionals and conditional probabilities. *Philosophical Review* 85: 297–315.
Lindworsky, J. 1916. *Das schlußfolgernde Denken. Experimentell-psychologische Untersuchungen*. Freiburg im Breisgau: Herdersche Verlagshandlung.
Lipton, P. 2004. *Inference to the best explanation*, 2nd edn. London: Routledge.
Lombrozo, T. 2007. Simplicity and probability in causal explanation. *Cognitive Psychology* 55: 232–257.
May, J., W. Sinnott-Armstrong, J.G. Hull, and A. Zimmerman. 2010. Relevant alternatives, and knowledge attributions: An empirical study. *Review of Philosophy and Psychology* 1: 265–273.
McCarthy, J. 1977. Epistemological problems of artificial intelligence. In *IJCAI'77: Proceedings of the 5th international joint conference on artificial intelligence*, 1038–1044. San Francisco: Morgan Kaufmann Publishers Inc.
McGee, V. 1989. Conditional probabilities and compounds of conditionals. *Philosophical Review* 98: 485–541.
Oaksford, M., and N. Chater. 1991. Against logicist cognitive science. *Mind & Language* 6: 1–38.
Oaksford, M., and N. Chater. 1994. A rational analysis of the selection task as optimal data selection. *Psychological Review* 101: 608–631.
Oaksford, M., and N. Chater. 2007. *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Oaksford, M., and N. Chater. 2009. Précis of Bayesian rationality: The probabilistic approach to human reasoning. *Behavioral and Brain Sciences* 32: 69–120.
Oberauer, K., and O. Wilhelm. 2003. The meaning(s) of conditionals: Conditional probabilities, mental models and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29: 680–693.
Over, D.E. 2009. New paradigm psychology of reasoning. *Thinking and Reasoning* 15: 431–438.
Over, D.E., I. Douven, and S. Verbrugge. 2013. Scope ambiguities and conditionals. *Thinking and Reasoning*. doi:10.1080/13546783.2013.810172.
Over, D.E., and J. St B.T. Evans. 2003. The probability of conditionals: The psychological evidence. *Mind & Language* 18: 340–358.
Over, D.E., C. Hadjichristidis, J. St B.T. Evans, S.J. Handley, and S. Sloman. 2007. The probability of causal conditionals. *Cognitive Psychology* 54: 62–97.
Pelletier, F.J., and R. Elio. 1997. What should default reasoning be, by default? *Computational Intelligence* 13: 165–187.
Pfeifer, N. 2012. Experiments on Aristotle's thesis: Towards an experimental philosophy of conditionals. *The Monist* 95: 223–240.
Pfeifer, N. 2013a. Reasoning about uncertain conditionals. *Studia Logica*. In press. doi:10.1007/s11225-013-9505-4.
Pfeifer, N. 2013b. The new psychology of reasoning: A mental probability logical perspective. *Thinking and Reasoning.* In press doi:10.1080/13546783.2013.838189.
Pfeifer, N., and G.D. Kleiter 2003. Nonmonotonicity and human probabilistic reasoning. In *Proceedings of the 6th workshop on uncertainty processing*, 221–234. Hejnice. September 24–27th, 2003.
Pfeifer, N., and G.D. Kleiter. 2005. Coherence and nonmonotonicity in human reasoning. *Synthese* 146: 93–109.
Pfeifer, N., and G.D. Kleiter 2006. Is human reasoning about nonmonotonic conditionals probabilistically coherent? In: *Proceedings of the 7th workshop on uncertainty processing*, 138–150. Mikulov. September 16–20th, 2006.
Pfeifer, N., and G.D. Kleiter. 2009. Framing human inference by coherence based probability logic. *Journal of Applied Logic* 7: 206–217.
Pfeifer, N., and G.D. Kleiter 2010. The conditional in mental probability logic. In *Cognition and conditionals: Probability and logic in human thought*, eds. M. Oaksford, N. Chater, 153–173. Oxford: Oxford University Press.
Pfeifer, N., and G.D. Kleiter 2011. Uncertain deductive reasoning. In *The science of reason: A Festschrift for Jonathan St. B.T. Evans*, eds. K. Manktelow, D.E. Over, S. Elqayam, 145–166. Hove: Psychology Press.
Polanyi, M. 1976. *The tacit dimension*. New York: Anchor Books.
Politzer, G., and G. Bourmaud. 2002. Deductive reasoning from uncertain conditionals. *British Journal of Psychology* 93: 345–381.
Politzer, G. 2005. Uncertainty and the suppression of inferences. *Thinking & Reasoning* 11(1): 5–33.
Quine, W.V.O. 1969. Epistemology naturalized. In *Ontological relativity and other essays*, ed. W.V.O. Quine, 69–90. New York: Columbia University Press.
Rips, L.J. 1994. *The psychology of proof: Deductive reasoning in human thinking*. Cambridge: MIT Press.
Schurz, G. 2005. Non-monotonic reasoning from an evolution-theoretic perspective: Ontic, logical and cognitive foundations. *Synthese* 146: 37–51.
Skyrms, B. 1980. *Causal necessity*. New Haven: Yale University Press.
Sloman, S. 2005. *Causal models: How people think about the world and its alternatives*. New York: Oxford University Press.
Stenning, K., and M. van Lambalgen. 2005. Semantic interpretation as computation in nonmonotonic logic: The real meaning of the suppression task. *Cognitive Science* 29: 919–960.
Stenning, K., and M. van Lambalgen. 2008. *Human reasoning and cognitive science*. Cambridge: The MIT Press.
Störring, G. 1908. Experimentelle Untersuchungen zu einfachen Schlußprozessen. *Archiv für die Gesamte Psychologie* 11: 1–127.
van Fraassen, B.C. 1989. *Laws and symmetry*. Oxford: Clarendon Press.
Waldmann, M., and Y. Hagmayer. 2005. Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31: 216–227.
Weisberg, J. 2009. Locating IBE in the Bayesian framework. *Synthese* 167: 125–143.