

BELIEF IN PSYONTOLOGY

Jonathan Weisberg
University of Toronto

ABSTRACT. Which is more fundamental, full belief or partial belief? I argue that neither is, ontologically speaking. A survey of some relevant cognitive psychology supports a dualist ontology instead. Beliefs come in two kinds, categorical and graded, with neither kind more fundamental than the other. In particular, the graded kind is no more fundamental. When we discuss belief in on/off terms, we are not speaking coarsely or informally about states that are ultimately credal.

How are full and partial belief related in psychology's ontology? Credence-first philosophers think partial belief is more fundamental. Lockeans, for example, say that to believe P is just to have high credence in P . Categorical-first philosophers on the other hand make full beliefs fundamental instead. Having credence x in P might just amount to having a categorical belief that P 's probability is x , for example.

Work in cognitive psychology supports a different view, however. In humans, beliefs come in both coarse and fine kinds, with neither more fundamental than the other. Epistemologists who focus on one kind to the exclusion of the other, or who treat one as central and the other as an afterthought, risk toiling at a fiction.

This conclusion is necessarily tentative. For one thing, the empirical work is nascent and ongoing, its results subject to revision. But more than that, it's sometimes unclear what the present results tell us about the ontological and epistemological questions of interest to philosophers. Still, a *prima facie* case can be made. And making it is an essential step in bringing our best science to bear on what is, in part, an empirical question.

Thanks to Kristen Aspevig, Imogen Dickie, Cian Dorr, Branden Fitelson, Hannes Leitgeb, Jennifer Nagel, Mark Schroeder, Adam Sennet, Julia Staffel, Scott Sturgeon, David Wallace, two anonymous reviewers for *Philosophers' Imprint*, and audiences and seminar participants at Rutgers University, the University of California at Berkeley, the University of California at Davis, the University of Geneva, the University of Leeds, the University of Oxford, and the University of Toronto for their generous and helpful feedback. I am especially grateful to Jennifer Nagel, whose work is the foundation for many of the ideas in this paper, and whose generous help with this project stretches over many years, and many patient conversations.

1. BACKGROUND

1.1. **Why Monism?** Why think one kind of belief is more fundamental than the other? It's not just that Anglophone philosophers use 'belief' for both kinds, so they must really be the same thing deep down.¹ It's rather what causes us to use 'belief' for both attitudes. They are supposed to do the same job. In folk psychology, full beliefs are supposed to guide inference, action, and assertion. Someone who believes P will infer conclusions from P , act on the assumption that P , and assert P , in appropriate circumstances (Ross & Schroeder, 2014; Schwitzgebel, 2015). But the theory of subjective probability assigns these roles to partial belief instead. Inferences are governed by one's prior degrees of belief via the rule of conditionalization (Hacking, 1967; Lewis, 2010); actions are similarly driven via expected utility maximization (Savage, 1954; Jeffrey, 1983); and assertions can be treated as a special case of action (cf. Brown, 2011, 2012; Douven, 2006, 2009).

Moreover, the theory of partial belief is supposed to be complete. Conditionalization determines everything you should conclude from the evidence given your prior degrees of belief. No other variables seem to factor in, not even your full beliefs. Expected utility maximization then determines what you should do and say based on those degrees of belief, given your utilities. Even when multiple acts have maximal expected utility, it is determined that you are permitted to choose any of those maximal acts. So there doesn't seem to be any room for full beliefs to make a difference.

A clash thus looms. Call it *the dualist dilemma*: if full beliefs are metaphysically distinct from partial ones, do they also govern inference, action, and assertion? If so, they either dictate the same inferences, actions, and assertions as one's partial beliefs, or they do not. If they do, then they appear to be extraneous, a theoretical extravagance. If they do not, then one's full beliefs can dictate a different course than one's partial beliefs, and our theory makes conflicting recommendations. It says an agent should draw a certain conclusion or take a certain action based on her full beliefs, and yet she also should not take it given her partial beliefs.² (Kaplan, 1996; Frankish, 2009; Stalnaker, 1984; Sturgeon, 2008)

Some well-known philosophical puzzles and problems arise from this clash. The preface paradox is a notorious problem for inferences based on full beliefs (Makinson,

¹Even in English it's often more natural to describe partial beliefs using terms like 'certainty', 'confidence', and 'credence'. In fact full beliefs may be more commonly attributed using *thinks that* than *believes that*. The verb *to think* ranks 12th among the most common verbs in American English where *to believe* ranks 50th (Davies & Gardner, 2010, 317). (Thanks to Jennifer Nagel for this datum.)

²There is a closely related, descriptive version of this dilemma, which afflicts our psychological and linguistic theorizing: the agent both will and will not so infer/act/assert.

1965). Given my full beliefs in the individual claims in my book, I can conclude that the book contains no errors. Yet my degrees of belief belie the absurdity of that conclusion: each individual claim is highly certain, but the conjunction of all of them is highly dubious. The bootstrapping problem (Fumerton, 1995; Vogel, 2000) can be seen as a related puzzle. Extending the inferential chain of the preface paradox, I can conclude that I am an extraordinarily reliable researcher. I've just written the first ever error-free book in my field! (Christensen, 2004; Weisberg, 2010, 2012)

In the realm of action and assertion, Bayesians criticize knowledge-based norms by citing conflicts between what your degrees of belief recommend vs. your full beliefs/knowledge. Suppose you know that Ada Lovelace wrote the first computer program. Should you gamble your life on it in exchange for a penny? Should you assert it to a friend who faces a similar decision? Certainly not, many Bayesians will say.³ But you should take the bet if you're to act on your full beliefs, even just those that constitute knowledge.

So the dualist's dilemma is this. Either full belief plays a substantive role in our psychology and epistemology, or it does not. If it does, then the role it plays is the fool. It leads to absurd inferences and irresponsible actions and assertions. If, on the other hand, it does not have a substantive role, then its appearance in a theory is pointless, even mysterious. In what sense is it something separate from partial belief?

Considerations of parsimony put additional pressure on a dualist ontology. A theory with just one kind of belief carries less ontological baggage in an obvious sense: it posits one fewer kind of thing. But dualism also threatens to posit twice as much of something else, namely *storage*. On a "belief box" picture, for example, there would have to be two boxes instead of one—one for partial beliefs and a second for full beliefs. That looks like a substantial extra storage burden. Yet humans' capacity for storing partial beliefs is a feat some philosophers already find hard to accept (Pollock, 2006, ch. 6).

Worse yet, how do dualists explain the tight connection between full belief and high credence? Typically, we believe P only when we are highly confident in P .⁴ And high confidence in P is often accompanied by full belief that P .⁵ Credence first philosophers have a built-in explanation for this correlation: full belief just is

³Hawthorne & Stanley (2008) report encountering this objection frequently. Cresto (2010); Douven (2008); Schiffer (2007) raise further Bayesian objections.

⁴But cf. §6.2.4 of Maher (1993).

⁵The lottery paradox (Kyburg, 1961) being the classic exception, generalized by Harman (1973) and Vogel (1990).

high confidence, or high confidence plus some further condition like “stability”.⁶ But dualists must, it seems, posit some sort of synchronization mechanism to keep the two kinds of belief roughly in line (or else pray for pre-established harmony).⁷ So on top of the extra storage burden comes the additional maintenance burden of synchronization.

We dualists have our work cut out for us then. We face two broad challenges:

- (1) *The Dualist Dilemma*. How can full belief guide inference/assertion/action, without either (a) resulting in absurd outcomes, or (b) becoming redundant with partial belief?
- (2) *The Parsimony Problem*. How are full and partial beliefs stored and maintained without duplicating cognitive burdens like storage and maintenance?

No wonder then that many philosophers prefer monistic alternatives.⁸

The temptation to embrace credence-first monism can be especially strong for those immersed in the Bayesian epistemological tradition, where partial belief reigns. Categorical-first monism certainly has its defenders (e.g. Lance, 1995; Easwaran, 2015), and much of my argument will be relevant to monisms that begin with full belief. But for simplicity, I will focus on credence-first monism. I suspect most of the resistance to dualism comes from credence-first epistemology anyway. And my main concern is to dislodge the view that credence should come first in epistemology because it comes first in psychology’s ontology.

1.2. Why Not Monism? *Eliminativism* is a radical form of credence-first monism which denies that full beliefs exist at all. Jeffrey (1970) famously held that full belief is a sort of legacy notion. It is a muddled leftover from naive folk psychology, now eclipsed by the superior, graded conception of belief established in the theory of subjective probability.⁹ With full belief excised from our psychology, our epistemology becomes

⁶Leitgeb (2013, 2014) endorses stability as a condition on rational belief, but not, importantly, as part of an ontological/reductive account. Leitgeb is explicitly agnostic on our ontological question.

⁷Lin & Kelly (2012) offer a formal theory that ensures pre-established harmony in a way. If credences are updated by conditionalization, and full beliefs are updated according to a method from Shoham (1987), then full and partial belief stay in sync. Importantly though, the appropriate Shoham-revision on a new piece of information doesn’t just depend on your existing full beliefs. It is determined by a plausibility ordering that aligns with the relevant credence function in a certain way. So synchronization is achieved by, effectively, consulting one’s partial beliefs every time one updates one’s full beliefs.

⁸Julia Staffel notes in personal communication that related issues can emerge even on a monistic view. Categorical-first monism can generate clashes too, between a belief that *P* and a belief that *P* is highly probable but not certain, for example. This may be another reason monists tend to be credence-first monists (see next paragraph). Full belief itself seems to be problematic, a tool too crude to be worth a serious theorist’s trouble.

⁹Ironically, Jeffrey seems to fall back on full belief to formulate his influential decision theory on p. 1 of *The Logic of Decision*. Which acts should an agent contemplate in framing a decision problem? The

stunningly simple: probabilism and conditionalization complete the theory. With expected utilities added in, we get an elegant decision theory into the bargain.

Such extreme simplicity doesn't come for free, of course. First, eliminativism is hard to square with introspection. Deliberation sometimes seems to end in a kind of inner assenting to *P*, or acceptance of *P*. And some snap-judgments don't have any apparent gradation in confidence attached. Take the judgment that my elbow itches, or that there is a taxi parked outside my window. I can generate confidence-levels for these judgments with a bit of contemplation. I might enumerate all the possible ways they could be in error, for example. But these confidence levels seem to be generated upon reflection, rather than searched out through introspection. After all they seem to fluctuate as I contemplate various possibilities of error. Second, eliminativism threatens a mass-error theory for folk discourse. All our attributions of categorical belief would seem to be false, according to eliminativism. Or maybe they are just so confused as to be meaningless.

Reductionism is less radical, alleviating some of these concerns. Reductionists don't deny the reality of full beliefs. They just make them ontologically secondary. Full beliefs reduce to partial beliefs, or at least supervene on them.¹⁰ The arch-reductionist is the Lockean, who holds that to believe *P* is just to have a credence in *P* above some threshold (Foley, 1993, 2009; Sturgeon, 2008). Some Lockeans will say the relevant threshold varies depending on what problems the agent faces, or what her cognitive values are (Christensen, 2004; Hawthorne & Bovens, 1999; Hawthorne, 2009). And other reductionists might prefer a more sophisticated criterion than a mere threshold (cf. Leitgeb, 2013, 2014). But the general idea is that full belief talk is a sort of abstraction over a reality that is fundamentally graded.

Reductionists avoid the mass-error theory threatened by eliminativism. On their view, ascriptions of full belief are often true. Just as dogs can be small, large, or middling, belief-states can either dissent, assent, or suspend (Christensen, 2004). It's just that the underlying reality is continuous: a dog's length can be anywhere from a few inches to several feet, and a belief's strength can vary continuously too.

Still, reductionism isn't entirely at peace with certain folk truisms and introspective data. For example, how would a Lockean make sense of inner assent or acceptance of *P* at the end of deliberation? She might just reject the alleged introspective data, as Christensen (2004, 98) does. But she might try to recover it instead. For example, she

ones he believes are available to him: "In the simplest cases the number of possible acts that the agent believes are available to him is finite [...]" (Jeffrey, 1983).

¹⁰'Supervenientism' might be a better title, but the difference between reductionist and merely supervenientist views won't matter for our purposes.

might say it's a cognitive act, a decision to stop deliberating based on the expected utility of continuing vs. stopping.

But if it is a cognitive act, that act still seems to leave us in a certain kind of state, something more than just being highly confident. Two characteristics distinguish this state. First, we become disposed to rely on P —to use it as a premise in future reasoning, to assume it in decision-making, and to assert it. These dispositions aren't exceptionless. If the implications of P are surprising, or if the stakes are unusually high, we may not take P for granted. But we otherwise will, in a way that we wouldn't have before we accepted or assented to it. Second, we become resistant to re-opening deliberation—we treat the question whether P as settled. We don't become *completely* opposed to re-opening deliberation. But it would require a solid cognitive nudge: substantial new counter-evidence, unusually high stakes, or similar.

Reductionists may demur here, denying that such a state ever results when we inwardly accept or assent to P . We will return to this challenge later in §§2–3. For now, let's run with the idea.¹¹ So what if we do arrive at such a state once we inwardly assent to P ? So what if we become disposed to use P as a premise and to resist re-opening deliberation about whether P ?

An instructive a priori argument against reductionism emerges. It begins with the observation that full and partial belief have essentially different functional roles.¹² Both are meant to govern inference, action, and assertion. But they do so in different ways. Partial beliefs function as *weights* in rules whose parameters are *continuous*. Bayes' theorem weighs prior probabilities against likelihoods, and the expected utility formula weighs probabilities against utilities. But full beliefs function categorically: their contents are taken as *givens* in procedures whose states are often *discrete*, like inference to the best explanation or dominance reasoning (Ross & Schroeder, 2014, §§1.3–1.4).

Now imagine a being inspired by the work of Levi (1967, 1980), *the Leviathan*. The Leviathan's mind contains two data structures.¹³ The first structure is an old-fashioned belief box, a set of propositions inscribed in mentalese. The second records a probability function, a map from propositions to real numbers. These structures are updated as follows. First, the probability function is updated in response to new evidence by conditionalization. Then the belief box is updated by calculating the

¹¹Even Lockeanism's chief proponent, Foley (1993, 2009), affirms that we become disposed to use P as a premise once we fully believe it.

¹²This observation and the resulting argument are drawn from (Ross & Schroeder, 2014, §§1.3–1.4). They deploy it in a slightly different context, and via an argument whose details differ. But I take the essential idea to be theirs.

¹³The Leviathan is so called because of its size.

expected cognitive value of accepting P .¹⁴ The Leviathan adds P to the belief box if doing so maximizes expected cognitive value. When a proposition is added to the belief box, its consequences are deduced and added too (including its consequences given the other contents). The probabilities are also affected, by conditionalizing on P . Finally, practical decisions are made as follows. When the propositions in the belief box identify an option as dominating all others, that option is taken. Otherwise, available options are evaluated by calculating expected utilities.

The Leviathan's probability assignments play the degree-of-belief role perfectly; they are weighed in conditionalization and expected utility assessments. The inscriptions in its belief box play the role of full beliefs; they are used as premises for drawing further conclusions, and for identifying optimal actions. So the Leviathan's full beliefs are not reducible to its partial beliefs, nor are they realized by those partial beliefs. Even a 99.9% probability for P would not constitute a full belief in this creature. For it will not (directly) cause the creature to assume P in any piece of reasoning, whether theoretical or practical. That 99.9% probability can only be weighed against the 0.1% probability of $\neg P$ to decide whether to accept P . If P is then accepted—and only then—will P be taken for granted, and further conclusions drawn from it.

1.3. Why Go Empirical? And yet, these a priori observations about the Leviathan do not settle our central question. They show that there is no necessary reduction of full belief to partial belief. There are possible creatures in whom facts about full belief are not determined by the facts about partial belief. But there may also be possible creatures in whom full beliefs nevertheless are realized by partial beliefs.

Imagine a different being now, *Rudy the Robot*, this time inspired by the work of Carnap (1971). Rudy's mind contains a probability map, like the Leviathan's. But unlike the Leviathan, there is no old-fashioned belief box. Rudy's doxastic plumbing is exhausted by the probability map, which is updated only by conditionalizing on new evidence. Practical decisions are made by calculating expected utilities.

Does Rudy have any full beliefs? I think he might. If Martians invade tomorrow and they possess minds like Rudy's, I might comfortably describe them in full belief terms.¹⁵ "They acted friendly at first because they knew it would make humans vulnerable"; "They attacked Las Vegas because their leader thought it was a military hub"; etc. Just because their brains don't have anything that plays the functional role

¹⁴The expected cognitive value of an act is formally the same as its ordinary expected value, $\sum_i p_i u_i$, except that cognitive utility takes the place of overall utility. That is, the utility function u only reflects values like truth, informativeness, understanding, etc. Exactly which values should be included varies by author.

¹⁵I'm grateful to Cian Dorr here.

of full belief perfectly doesn't mean they don't have anything that comes close. In their case, strong partial beliefs might come closest to playing that role, even close enough to ground full belief talk (cf. Lewis, 1970). So in some cases, full beliefs might be realized by strong partial beliefs, as the Lockeans say.

We come then to an empirical question: what is the architecture of the human mind, and what in it plays the role of full belief best (if anything)? Are we like the Leviathan, with two separate data structures, one of which has the stronger claim to 'full belief', the other to 'partial belief'? Or are we like Rudy the Robot, where the mechanisms that merit the name 'partial belief' also have the best claim to 'full belief'? Or is neither picture apt in our case? As Leitgeb (2013) says, when it comes to ontological questions about full and partial belief, "one will need to carry out some empirical investigations into belief along the lines of: what are the agent's belief-generating systems like?"

We will draw on three literatures from cognitive psychology to begin answering these questions. Our tentative finding will be that neither picture is entirely apt—we are not built like either Rudy or the Leviathan. But we are more like the Leviathan in one important respect: our full and partial beliefs are realized in largely separate mechanisms. In us, full and partial belief are largely metaphysically distinct. They do share a common supervenience base to some extent. But not enough to merit treating either as secondary, in either our psychological or our epistemological theorizing.

2. JUDGMENT & DECISION-MAKING

How do our beliefs function in reasoning? Do we weigh partial beliefs using rules like Bayes' theorem and expected utility maximization? Do we instead rely on full beliefs for premises to be used in natural deductions, enumerative inductions, and inferences to the best explanation? Or do we do something entirely different, perhaps something that employs both full and partial beliefs simultaneously? To find out, we must look at the research on human judgment and decision-making.

2.1. Empirical Background. Famously, psychologists began to reject the classical model from economics in the 1970s. Humans did not seem to be perfect Bayesian agents whose beliefs obeyed the probability axioms, were updated by conditionalization, and were weighed to maximize expected utility.

Kahneman & Tversky (1973) found subjects persistently ignoring base-rates, for example. When given a description of a person, Tom, and asked to judge the probability of his being an engineer, they ignored information about the frequency of engineers. They seemed to consider only how representative Tom was of the class of engineers—how similar he was to the typical engineer. Kahneman & Tversky

hypothesized that people rely on such similarity judgments as a kind of shortcut, a heuristic they dubbed *representativeness*. Tversky & Kahneman (1983) later found the majority of their subjects committing the conjunction fallacy in ways one would expect if they were using the representativeness heuristic.

Tversky & Kahneman (1973) posited another heuristic, availability. Where does the letter R appear more often in English words: as the first letter or as the third? Most subjects said it was more likely to appear in the first position, though the third position is actually more common. The availability heuristic explains this pattern. It's easier to think of words that begin with R than to think of words where it appears in third position. Since examples of the first position are more readily available, they are judged more common.

Classical assumptions about decision-making were also found to have reliable violations, and heuristics explaining them were posited. For example, Tversky (1972) studied the method of Elimination By Aspects (EBA): to choose a flight from a list of options, first eliminate the expensive ones, then the ones with numerous layovers, and finally those on airlines you dislike, until just one remains. In general, EBA considers the desirable features an option might have, starting with the most important features first. It then eliminates options lacking each feature until one option remains.

EBA is appealingly simple, but it can fail to maximize utility. An option that doesn't make the first cut isn't considered anymore, even though it may have enough other desirable features to make it optimal.¹⁶ For example, you might eliminate the only direct flight because it costs much more than you initially expected to pay. But had you realized ahead of time just how bad the remaining flights would turn out to be (numerous layovers, ruthless airlines), you would have been willing to reconsider paying more for a direct flight.¹⁷

As more and more heuristics emerged, the "adaptive toolbox" paradigm came to dominate in the 1980s and 1990s. On this approach, humans form judgments and make decisions using a variety of methods, choosing our methods adaptively for the circumstances. For example, in the effort-accuracy framework of Payne et al. (1993), we choose the tool that best balances the competing demands of accuracy and effort. That is, we generally use the decision rule that takes the least effort to use while still

¹⁶Tversky (1972) was originally motivated by different considerations, like the following example from Luce & Raiffa (1957). Suppose you can't decide between a trip to Paris and a trip to Rome: $Paris \sim Rome$. Now a third option comes along, $Paris^+$, a trip to Paris plus a \$1 bonus. Naturally, you prefer $Paris^+$ to Paris: $Paris^+ > Paris$. But you might still be torn between $Paris^+$ and Rome: $Paris^+ \sim Rome$. So $Rome \sim Paris^+ > Paris \sim Rome$, contra expected utility theory.

¹⁷Indeed, many people would go back and reconsider. This shows that we can monitor our use of heuristics like EBA. We sometimes correct heuristic reasoning when we notice it underperforming (Stanovich, 2011).

performing accurately enough in the circumstances. If you're pressed for time and all the available flights are similar, you might use a quick-and-dirty method like EBA to pick one. If instead you're deciding national vaccination policy, with many lives at stake and enough time to do intensive calculations, you might use expected utility maximization instead. (See Gigerenzer et al. 1999 and Payne & Bettman 2004 for surveys and more details.)

More recently, the Evidence Accumulation Model (EAM) has been proposed in the hopes of unifying the disparate tools in the box (Lee & Cummins, 2004; Newell et al., 2007; Newell & Lee, 2011). According to EAM, we always decide using the same method. We examine available evidence, starting with the most significant evidence first, and evaluating the support each piece of evidence lends to each option as we go. The process stops when one option emerges as sufficiently well supported.

Take the question which flight will be best, all things considered. Relevant evidence includes things like price, number of stops, airline reputation, etc. Assuming price is the best predictor of which flight is best, we start with price. If one flight is much cheaper than the others, the evidence might already be sufficient to settle the question, and EAM can stop there. But if not, it'll go on to consider number of stops. Maybe now one flight emerges as clearly superior—it's almost the cheapest, and it's the only direct one. Or maybe not, in which case EAM goes on to consider airline reputation, etc.

EAM applies to judgments as well as decisions.¹⁸ Suppose you're asked which city is larger, San Diego or San Antonio.¹⁹ You might begin by considering which city is more familiar. If that still leaves some doubt, you might go on to consider how many major sports teams from each city you can name. Then you might ask whether either is a state capital. And so on, until one answer emerges as sufficiently supported by the considered evidence.

If EAM always applies the same procedure, how does it account for variation? Why do we sometimes evaluate our options in depth, other times only superficially? The model contains an adjustable parameter, what counts as "sufficiently" well supported by the evidence. EAM terminates when one option exceeds a pre-determined threshold of evidential support. A lower threshold will generally result in shallower evaluation of the evidence, a higher threshold tends to require more evidence to be considered. The threshold can be set depending on availability and need. When we need to be sure of our answer's correctness, the threshold can be set higher. When

¹⁸So do many of the methods posited by the adaptive toolbox approach. Indeed, psychologists often treat the boundary between practical and theoretical reasoning less rigidly than philosophers do. The same or similar mechanisms are often taken to drive both judgment and decision-making.

¹⁹The example is from Goldstein & Gigerenzer (2002), who study it in a different context.

resources like time and attention are limited, it can be set lower. The adaptive toolbox is thus replaced by a single, adaptable tool.

2.2. Philosophical Application. Categorical judgments appear in all these models. The representativeness heuristic uses assumptions about Tom to judge how representative he is of the class of engineers. The availability heuristic uses assumptions about where the letter R appears in certain English words to estimate how commonly it appears first rather than third. Elimination by Aspects (EBA) chooses a flight based on assumptions about how much each costs, how many stops each makes, etc. And the Evidence Accumulation Model (EAM) does the same.

Credence-first philosophers may respond that these categorical judgments are really graded attitudes at bottom. I see two ways of trying to make this response work.

The first grants that methods like EBA and EAM function as advertised. They genuinely do rely on categorical assumptions. But these assumptions are extracted from graded beliefs. The subject has a very high credence that the flight costs \$500, for example. And that credence is so high as to be practically 1 for present purposes. So the possibility that the price is actually something else is set aside, and the proposition *The flight costs \$500* is delivered to the relevant decision-method to be used as a categorical assumption.

There are reasons to be skeptical of this proposal. First, it's less parsimonious. It requires two extra computational steps for each assumption used. Before an assumption can be made, its degree of belief has to be compared to some threshold, and the possibility that it's false must then be discarded.²⁰ Second, these are exactly the kinds of extra computational operations an adaptive decision maker is supposed to minimize. Many researchers decompose heuristics into just such computational steps in order to estimate the time and effort each heuristic costs (Bettman et al., 1990; Huber, 1980; Newell & Simon, 1979; Payne et al., 1993). Indeed, Bettman et al. estimate that discarding a possibility is by far the most expensive operation in terms of effort—almost twice as expensive as the next-most-effortful operation, and between two and eight times as effortful as all others. In terms of time, it is the second most expensive operation by their estimate, with the others again coming in well behind. Figure 1 displays the relevant table of results.

But even if this proposal were true, it would give away the game. For it grants that there is a state, separate from a credal one, that functions like a full belief. Even if a

²⁰Other computational operations may be necessary too. Would-be assumptions that are tested against the threshold and fail still cost a comparison step. And the threshold has to be set somehow initially.

	Response time	Self-reports of effort
READS	1.19*	.10*
ADDITIONS	.84*	.08*
PRODUCTS	2.23*	.19*
COMPARISONS	.09	.04
ELIMINATIONS	1.80*	.32*
DIFFERENCES	.32	.12

* Significantly different from zero at $p < .05$.

FIGURE 1. Adaptation of Table 7 from (Bettman et al., 1990): Coefficient Estimates of Response Time and Self-Reports of Effort for Elementary Information Processes.

process like EBA draws on credal states to supply the contents it uses as assumptions, the extracted contents are then deployed as categorical assumptions. The existence of such a process is enough to disqualify high credences as the realizers of full beliefs. They may causally affect which propositions are fully believed, but the full belief that results as an effect lives, at least partly, in the computational states of the process that uses the proposition as a categorical assumption.

Here's another way to see the point. Recall what tempts the credence-first theorist: the dualist's dilemma. Taking full belief seriously leads to nasty puzzles like the preface paradox, but those problems dissolve when we make partial belief primary. Then Bayesian epistemology and decision theory provide a satisfactory and complete normative theory. The puzzles of full belief can be written off as symptoms of loose talk and crude thinking about a reality that is fundamentally fine. But now we're admitting that categorical assumptions appear in the processes by which we judge and decide. So we're right back with the same problem. If you're entitled to assume that the first claim in your book is true even though it's only 99.9% certain, you're entitled to do the same with all the others. But then you're entitled to conclude they're all true, and to assume so in deciding whether to stake your life on your book's perfect accuracy in exchange for a penny. In other words, we're right back with the problem of reigning full beliefs in, to keep them from generating foolish conclusions and decisions that our theory of partial belief easily avoids.

So what's the second way of insisting that methods like EBA and EAM ultimately rely on graded attitudes? It's to deny that these methods actually operate as advertised. They don't really use categorical assumptions, because these "assumptions" aren't treated as givens. Rather, they are weighted according to their respective credences. In EAM, for example, each item of evidence is given weight in proportion to its credence. If the fact that a flight is direct supports its superiority to degree 0.5, but

you're only 0.99 sure it's actually a direct flight, then the support is only counted at 0.495. And similarly for each other piece of evidence accumulated.

Other methods are harder to render in such weighty terms. Take EBA, which is thoroughly discrete. At each step an option is either eliminated for lacking the desirable feature in question, or not. If it is eliminated, it is out of consideration. It won't be examined at any later step. Unlike EAM, which accumulates quantities of evidential support, EBA tracks no quantity to which a credal weight can be applied. This is essential to the cognitive savings EBA affords. Eliminating options from consideration saves the work of continuing to assess them. Other methods in the adaptive toolbox are similar in this regard, like the lexicographic rule (Coombs et al., 1954; Tversky, 1969; Payne et al., 1993) and the take-the-best heuristic (Gigerenzer & Goldstein, 1996; Gigerenzer et al., 1999).

Could credence-first theorists pin their hopes on EAM winning the day? Maybe the adaptive toolbox paradigm, with its coarse methods like EBA, will prove too crude. Maybe EAM will prove the better model, a model that can be altered to use credal weights instead of categorical assumptions.

Maybe, but the proposed alteration would be much more drastic than first appears. Revisiting Figure 1, we find that `PRODUCT` is by far the most expensive computational operation in terms of time, and the second-most expensive in terms of effort. If EAM is modified so that evidence has to be weighted, the result is a much more expensive procedure than what these researchers are actually proposing.

In fact, this modification starts to resemble the most expensive method, the "weighted additive" rule, WADD. The WADD rule is formally equivalent to expected utility, and is thus regarded as only occasionally used, at most. The modified version of EAM would still have significant savings over WADD. It would only weigh and add evidence until the pre-determined threshold is crossed, whereas WADD weighs and adds all the available evidence. But the similarity still highlights what's being lost: significant cognitive savings.

And for what gain? The difference in weighted evidential support will be inverse to the subject's level of certainty. If she is 99% certain the flight is direct, this is the difference between 0.5 evidential support and 0.495, a difference of 0.005. Such differences would hardly ever affect the outcome of an EAM decision.

Categorical-first philosophers can't claim victory here either, though. We've seen several places for credal weights to operate within the models surveyed.

Most obviously, occasional uses of expected utility maximization are by no means ruled out. Departures from expected utility are where the novelty is, so the research focuses there. But it's generally allowed that we do sometimes calculate and maximize

expected utility, especially in cases where it's easy to do. And even when we don't calculate expected utilities, we may still use procedures that are structurally similar. At least in some cases, we might still weigh credences against utilities, just in biased or inaccurate ways. Kahneman & Tversky (1979)'s prospect theory is an influential model of this sort, and numerous extensions and developments of prospect theory have been proposed and continue to be explored. Of special interest to us are extensions to cases where probabilities are not explicitly given (Wu et al., 2004). In these cases, the probability-like weights are even more plausibly understood as degrees of confidence.

But even within the cognitively more economical models we focused on here, there is plenty of room for partial beliefs to do their weighty work. EBA, for example, considers aspects in order of "weight". An aspect's weight can be understood as its importance to the subject, the value she places on it. But it can instead be interpreted as a credence: the subject's credence that an option is best, given that it has the desirable feature in question.

A similar interpretation may be even more promising for EAM. Conditional probabilities determine the order in which EAM examines evidence. The quantity $p(A > B | F(A) \wedge \neg F(B))$ determines the priority given to the evidence that A has feature F and B does not, where $A > B$ means that A is a better option than B . These probabilities indirectly determine how much weight a piece of evidence is given, since items considered first have a greater chance of determining the outcome.²¹ But also, the amount of evidential support $F(A) \wedge \neg F(B)$ gives to $A > B$ is measured by the log-likelihood ratio,

$$\log \frac{p(F(A) \wedge \neg F(B) | A > B)}{p(F(A) \wedge \neg F(B) | B > A)}$$

So there is a second, more direct way in which these probabilities are used to weigh the evidence in EAM. Given the implicit role these probabilities play here, and the long Bayesian tradition of seeing degrees of belief as largely implicit attitudes exposed primarily through behaviour, it seems fair enough to interpret them as degrees of belief. Their functional role has the weighty, implicit character stereotypically associated with partial belief.

Dualism thus emerges as a strong candidate. The methods we use to form judgments and make decisions draw on both full beliefs and partial beliefs. In fact, we seem to draw on both kinds of attitude simultaneously. In EAM, for example, we rely on full beliefs for evidence, and we rely on partial beliefs to weigh the import of that evidence (Weisberg, 2013).

²¹A similar point may apply to the various methods EAM is meant to unite, like the lexicographic rule and the take-the-best heuristic.

This understanding also helps loosen the grip of the dualist's dilemma. If both full and partial belief govern inference/action/assertion, what keeps them from tripping over one another? Essentially, the answer is that no proposition is treated as both a given and a mere probability at the same time. At any given time, *P* is either taken as an assumption or else it is weighted according to its subjective probability, not both. Take EBA as an illustration. I assume the first flight costs \$500 as listed, the second \$600, and so on. Although there is a dispositional sense in which I am only 99% certain of these facts, that dispositional belief remains latent in ordinary circumstances. My partial beliefs figure in elsewhere in the decision process. I rely on my conditional credence that the cheapest flight will be better all-things-considered in order to decide which aspects to consider first. So at no point is there a proposition whose full and partial belief are active simultaneously.²²

When does my partial belief that the first flight costs \$500 have its day in the sun then? Many things might trigger a more cautious or scrutinizing treatment of the listed prices. We'll encounter some of these factors later in §4. But very briefly, if a lot is at stake and I have the time for a careful assessment, I might use expected utilities instead of EBA to pick my flight (Payne et al., 1993). Then the possibility that the listed flight is wrong will get the 1% weight its credence entitles it to. Or, if I idly wonder about all the ways in which a listed price can turn out to be wrong, I might shift from treating the list price as a given to treating it as just highly certain (Nagel,

²²An anonymous referee raises the following, natural question. Suppose your dispositional credence in *Q* is 0.9, while your dispositional credence in *Q* given *P* is 0.95. Suppose also that you fully believe *P*, though your credence in *P* isn't quite one (dispositionally again, for both attitudes). When your full belief in *P* becomes occurrent, your occurrent credence in *Q* will presumably be .95. But we said that your credence in *Q* was only .9. What should we make of this clash?

The puzzle exposes a gap between dispositional and occurrent attitudes. This gap is general, in several respects. It affects attitudes besides belief, like desire and fear. But also, we should acknowledge its presence in the case of belief quite apart from the debate over full and partial belief. Even on a credence-only view like eliminativism, one might be generally disposed to be .9 confident in *Q*, yet be .95 confident in a given instance. Even the most hard-core Bayesians will presumably admit that their occurrent confidence sometimes wavers.

We face interesting choices in such cases. We could say that your dispositional credence is actually indeterminate to some extent, contra the stipulations of the case, and that your occurrent credence falls within the bounds of that indeterminacy. Or, we might allow that the occurrent credence falls outside the bounds of your dispositional credence. I am sympathetic to the thought that both kinds of case can arise. And that invites the further, normative question whether your occurrent credence is rational in the second case, where it really does clash with your dispositional credence.

Assuming such cases can arise, my hunch is that they will be cases of bounded rationality, at least sometimes. Your occurrent credence clashes with your dispositional credence, and the dispositional one presumably reflects a more carefully considered point of view. There, the simplifying assumption that *P* is not being made. Assuming there is some sense in which it is rational for humans to make such simplifying assumptions, other judgments arising from them are plausibly rational in the same sense. But that kind of rationality may only be the bounded kind, since it reflects our need to make simplifying assumptions.

2011). Expecting to have to defend one's beliefs might also trigger a more careful assessment (Lerner & Tetlock, 1999; Nagel, 2010).

This kind of switching—between categorical judgment and more guarded, graded assessments—offers a promising tool for dualist resolutions of the preface paradox and other instances of the dualist dilemma. Nagel (2011) offers a diagnosis of the Harman-Vogel paradox in this vein. And Weisberg (2010, 2012) adapts Nagel's treatment while arguing that the Harman-Vogel paradox is of a kind with the preface paradox and the bootstrapping problem.²³ A detailed development of this approach to the preface paradox in particular would require a lengthy manuscript of its own, however.

What matters for present purposes is that the literature on judgment and decision-making surveyed here does the following:

- (1) It supports dualism directly, since the prominent models in the literature are highly amenable to a dualist interpretation.
- (2) It resolves the dualist dilemma in the abstract. These models show how full and partial belief can cooperate in effective and efficient reasoning, rather than stumbling over one another, or colliding with foolish results.
- (3) It opens up promising avenues for treating specific illustrations of the dualist dilemma, like the preface paradox.

One caveat is in order before moving on. We've looked at the dominant paradigms in judgment and decision-making research from the last few decades. But there are other research programs in this field, including one in particular that may appear less hospitable to full belief. Oaksford & Chater (1998, 2007, 2009) champion a Bayesian approach designed to explain the departures from classical models of rationality that began to accumulate in the 1960s and 1970s. For example, they offer a Bayesian explanation of the Wason selection task, where subjects seem to reliably violate basic rules of propositional logic (Wason, 1966). Oaksford & Chater are primarily concerned to show that what look like irrational violations of laws of deductive logic are actually rational attempts to make efficient and effective use of available information, by relying on probabilistic inference. Subjects in the Wason selection task are not failing to perform correct deductive reasoning. They are succeeding at using efficient probabilistic reasoning that is, usually, effective.

Given its focus, Oaksford & Chater's program isn't necessarily at odds with the other research we surveyed. They don't target the experimental findings that drive the work in the adaptive toolbox and EAM paradigms as explananda for their theory.

²³See also Dodd.

They focus instead on logical, “syllogistic” reasoning. Moreover, their approach may yet be hospitable to dualism. The Bayesian inferences and heuristics Oaksford & Chater posit rely on categorical inputs, much like EAM. They then work by weighing the probabilistic implications of those categorical assumptions—again, like EAM. So credence-first philosophers should be cautious about pinning their hopes on this program winning the day.

It bears reiterating, all the same, that the conclusions drawn here are at least as tentative as the empirical research supporting them. It’s early days for cognitive psychology, and for concrete models of judgment and decision in particular. Still, our best bet is to look to the best science we have so far.

3. COGNITIVE CLOSURE

The second feature that distinguishes full belief is its settled character. As we noted earlier (p. 5), introspection suggests a kind of assenting or accepting of P that ends deliberation about whether P . And once we’ve accepted P , we seem to become less willing to reopen deliberation about it.

Rudy the Robot’s high probabilities don’t seem to do anything that corresponds to these phenomena. Nothing distinctive happens in him when P ’s probability crosses a threshold of 90% or 95% that doesn’t also happen when it crosses a threshold of 5% or 10%. Rudy doesn’t inwardly affirm P when it crosses some high threshold. He doesn’t inscribe it in some special place in his heart. He just goes on giving it weight in proportion to its probability, the same as he would if it were only 10% probable. Nor does he become specially resistant to reconsidering P . He continues to re-evaluate P ’s probability the same as every other proposition’s, revising it constantly via conditionalization.²⁴

Leitgeb’s (2013; 2014) stability theory of belief might seem able to supply the desired resilience, purely in terms of Rudy’s partial beliefs. But on Leitgeb’s theory, “fully believed” propositions are evaluated and re-evaluated exactly the same as other propositions. They are only stable in the sense that evidence capable of dislodging them, i.e. of pulling them down below the threshold for full belief, is disbelieved. These disbelieved propositions are not in any way resisted. And Rudy’s degrees of

²⁴What about probability 1? Perfectly certain propositions may become specially resistant to revision, especially on the standard axiomatization of probability theory. But if Rudy’s full beliefs are only those that have credence 1, and can never be reconsidered as a result, he is a pathological sort of full believer. Acceptance is irrevocable for him. He might try to compensate by being very cautious about where to assign probability 1: only to evidential propositions about sense-data, for example. But then his full beliefs aren’t just bizarrely irrevocable, they are also few, far between, and much weaker than the sorts of full beliefs we customarily attribute. Even defenders of the view that full belief is credence one find such sparse and stubborn doxastic corpora unappealing (Dodd; cf. also Levi 1980).

belief in fully believed propositions still fluctuate the same as all others according to conditionalization, whether the new evidence was previously disbelieved or not. So there is no sense in which Rudy resists further inquiry into the truth of believed propositions.

So full belief's functional role includes an element—settledness—which partial beliefs do not capture. At least, partial belief is a poor fit on the classic Bayesian model. Is there anything in humans that captures this element better, and thus has a stronger claim to the name 'full belief'?

Nagel (2008, 2010) argues that the phenomenon of *cognitive closure* has such a claim, and I will follow suit.

3.1. Empirical Background. Cognitive closure is the closing of one's mind on a given question, "the juncture at which a belief crystallizes and turns from hesitant conjecture to a subjectively firm 'fact'" (Kruglanski & Webster, 1996, 266). This crystallization serves to bring deliberation and inquiry to an end, since "our time is finite, and hence the search [for relevant information] must be terminated at some point" (Kruglanski, 2004, 2). The sense of having established a firm fact also serves to prevent the potentially endless process of deliberation from constantly restarting. It delivers "a sense of secure knowledge that obviates our felt need for further agonizing deliberation" (p. 2), and enables us to "shut off our minds to further relevant information" (p. 4). Moreover, we are motivated to achieve and maintain this settled state by what Kruglanski (1989) dubs the *need for closure*, or NFC. This need has a phenomenological profile: "negative affect when closure is threatened or undermined and positive affect when it is facilitated or attained" (Kruglanski & Webster, 1996, 264). When NFC is strong, one can feel discomfort with ambiguity and a sense of urgency to resolve it, followed by a sense of relief once closure is achieved.

The idea of a settled cognitive state that we seek and maintain has a long history in psychology. Precursors can be found in the cognitive consistency theories of the 1950s, like Festinger's (1957) theory of cognitive dissonance or Heider's (1958) notion of cognitive balance. On these views, we are motivated to be in a state free of internal contradiction or tension. Kagan (1972) postulates a more general "wish to know" and eliminate uncertainty, but discovering internal inconsistency still plays a central role in Kagan's understanding of this motivation. The contemporary notion of cognitive closure, and the need for it, are different. We are motivated to be in a state free of ambiguity and uncertainty, not just of contradiction. And this motivation stems

from an evolutionary necessity, the need to achieve firm judgments we can use as a basis for action, rather than dithering endlessly.²⁵

That similar notions have been posited and studied by psychologists for decades provides a bit of evidence that something like cognitive closure exists. But some credence-first philosophers deny acquaintance with the phenomenon, like David Christensen:

I may start a picnic having heard a very positive weather forecast [...] But during the course of an hour, as clouds appear on the horizon and move toward us, as the sky gradually darkens, and as the breeze becomes stronger, my confidence in having a pleasant time fades, [...] at the end of the hour, I would unhesitatingly say “I believe that our picnic is going to be spoiled.” But at no point during the process do I seem to experience a discrete qualitative shift in my attitude toward the proposition that we’ll have a great picnic—no jumps from an inner “saying yes” to an inner withholding of judgment to an inner “saying no.” (Christensen, 2004, 98)

So let’s briefly look at more direct evidence, especially for the contemporary notion of cognitive closure specifically. What concrete predictions does the theory make and how have they been borne out?

People’s need for closure varies. Some people tend to have a strong need for closure, others less so. An individual’s baseline NFC is measured by the Need for Closure Scale, or NFCS, developed by Webster & Kruglanski (1994). They had 49 subjects fill out a questionnaire designed to probe their preference for order and structure, distaste for ambiguity, and desire for secure and stable knowledge. Three months later, the subjects took the questionnaire again. The results were very similar each time, suggesting that the questionnaire measures a stable personality characteristic. They then distributed the questionnaire to advanced classes of accounting majors and of studio-art majors, on the assumption that the accounting majors would tend to have higher NFC. As predicted, the accounting majors scored significantly higher, suggesting that the stable characteristic measured is in fact NFC.

In another study they found that subjects who scored high on the NFCS questionnaire were more influenced by earlier information than by later information. When given negative and then positive information about a job candidate, they were more likely to evaluate the candidate negatively than their low-scoring counterparts. When they were given the positive information first, they were more likely to give a

²⁵See chapter 4 of (Kruglanski, 2004) for a thorough survey of historical antecedents and their contrasts with the contemporary notion.

positive evaluation than their low-scoring counterparts. As expected, a higher need for closure corresponded to a tendency to be discount later information. Subjects who scored high on the questionnaire also requested less information on which to base their evaluation of a job candidate, expected to be more confident of their assessment, and took less time to make that assessment.

But an individual's NFC also varies with the circumstances. The motivation to settle a question can be intensified by increasing the costs of continued deliberation. If a subject is pressed for time because they need to make a decision that hangs on the question at hand, their need for closure will increase. Trivial factors can also increase NFC by making thinking more difficult, like being tired, intoxicated, or distracted by background noise. A number of studies verify that these intrapersonal variations in NFC have effects similar to those of the interpersonal variations measured by the NFCS (Kruglanski & Webster, 1991; Webster, 1993; Kruglanski et al., 1993; Webster et al., 1996). On the flip side, if the costs of being wrong are high or the subject expects to be held accountable for their answer, they will be motivated in the opposite direction: to extend deliberation instead (Kruglanski, 2004; Kruglanski & Chun, 2008).

So we have two ways of manipulating NFC in experiments and then testing predictions. We can measure interpersonal differences using the NFCS, and then see how people with different baseline levels of NFC behave. We can also manipulate NFC intrapersonally, by imposing a decision deadline, introducing background noise, or supplying alcohol.

Two broad effects are predicted by a high NFC: "seizing" and "freezing". With heightened NFC should come a tendency to seize on whatever information or hypothesis is offered in order to settle an open question. But there should also be a dual tendency, to freeze one's opinion about a closed question, becoming resistant to new information or hypotheses that might threaten that opinion. So we have two prediction schemas:

$$\begin{aligned} \text{Open Question} + \text{High NFC} &\longrightarrow \text{Seize} \\ \text{Closed Question} + \text{High NFC} &\longrightarrow \text{Freeze} \end{aligned}$$

In other words, when the subject's mind is open on a certain question, heightening her NFC should increase her tendency to seize on new information and hypotheses. But if her mind is already closed, heightening her NFC should incline her instead to resist new information and hypotheses.

Kruglanski et al. (1993) verified this pattern in a series of experiments where subjects acted as jurors. Everyone was given a description of the evidence in a legal

case. Some subjects were also given a legal analysis, which argued for a specific verdict on the basis of the evidence. The rest were given just the evidence. When subjects were given the legal analysis (closed question), those with higher NFC were more resistant to persuasion (freeze). Absent the legal analysis (open question), those with a higher NFC were more persuadable (seize). This pattern held whether NFC was varied intrapersonally via background noise (Experiment 2), or interpersonally as measured by the NFCS (Experiment 3). Kruglanski & Webster (1996) and (Kruglanski, 2004, ch. 5) survey numerous other results.

3.2. Philosophical Application. Interpreting the work on cognitive closure is straightforward for dualists: cognitive closure is the formation of a full belief. Full beliefs serve as premises in decision-making, and we are motivated to establish such beliefs when a decision is impending because they greatly simplify the decision-making process, as we learned from the literature on judgment and decision in §2. Once a full belief is established, we become reluctant to reopen deliberation because it is cognitively expensive. This reluctance is facilitated by the phenomenology of having settled a “firm fact”. The phenomenology of “crystallization”, where feelings of discomfort with uncertainty and urgency to resolve the question are replaced by relief, accompanies a categorical shift. As Nagel puts it, “Where high degrees of confidence are compatible with a continuing feeling of uncertainty, outright belief yields the subjective sense of a solid result.” (Nagel, 2010, 418) Confidence in *P* with some lingering doubt is replaced by an unquestioning view of the world as being such that *P*.

How might credence-first philosophers account for cognitive closure? Bayesians have long analyzed the search for evidence, and its termination, in terms of expected utility. So they will probably offer a similar analysis here.

Terminating deliberation is a decision, they may say. It is a cognitive choice based on a cost-benefit analysis. Deliberation takes time and effort, but it also leads to more informed decisions down the line. These costs and benefits are weighed according to one’s partial beliefs, as in any decision. We are motivated to seek information about questions on which we must act because getting more information has higher expected utility, provided the cost of acquiring it is negligible (Good, 1967). But when we achieve an extreme credence (either 1 or 0), the benefits of continued deliberation tend to become even more negligible. The chances of a significant change in opinion become too small. So we leave off deliberating, feeling relieved that we can turn our minds to other things. As for reopening deliberation, the situation is similar. Whether one should reconsider a question is to be decided on the basis of the potential costs

and benefits, weighted according to one's degrees of belief. The more certain one is that P , the less likely further deliberation is to make any difference. So having left off deliberating about P , one tends not to take it up again.

The trouble with this story is that it rests on a confusion. To succeed, it must confuse a state's cause for the state itself. Bayesians may be right that we decide when to close our minds on a question based on the expected utility of continuing inquiry. But the state that results—closed-mindedness about P —is not the same as the credal state that caused it.²⁶

One way to bring out this confusion is to ask why a decision against re-opening deliberation doesn't need to be made over and over again. At any moment following the initial decision to stop deliberating, the agent could resume deliberating. Does she repeatedly choose, moment-by-moment, against reopening deliberation? This would be an incredible expense of cognitive effort where there were supposed to be savings. Maybe instead the decision doesn't arise, because her mind has turned to other topics? But she will frequently turn her attention back to P , as she makes the P -dependent decisions that caused her to investigate P to begin with. A dualist will say that she doesn't reopen her inquiry about P because she has entered a new state, a state of seeing P as a closed matter. And closed matters aren't reopened without strong cause. This state may be brought about by a cognitive decision to stop deliberating. But it is an *effect* of that decision, rather than being *constituted by* that decision. Otherwise, nothing stops the decision from constantly re-arising.

A second way to bring out this point focuses on the "freezing" phenomenon. Subjects who close their minds on a question don't just stop deliberating and seeking evidence. They also become resistant to further evidence and hypotheses presented to them. They dismiss counter-evidence and discount dissenting opinions. According to the Bayesian story, this resistance is based on an expected utility calculation. Unless the evidence is particularly strong, the subject ignores it because the change in her opinion won't be large enough to make a difference to any future decision she is likely to encounter. In order to perform this calculation, however, she first needs to calculate what the change in her opinion would be. She only thinks the change isn't worth making because she knows it would be small. But if she's already calculated that much, she's already done the cognitive work she was supposed to be saving. By plugging the result into an expected utility calculation, she actually adds to her cognitive labours. (And if she then decides to ignore the evidence, she discards the results—a terrible cognitive waste!)

²⁶Special thanks to Jennifer Nagel and Adam Sennet for their generous help with the ideas in the following discussion.

A natural rejoinder is to say that she doesn't do the cost-benefit analysis directly. Rather, she uses a rough rule-of-thumb. When she decides to leave off deliberating, she adopts a policy of dismissing further evidence unless it's especially strong or striking. So she doesn't have to calculate the hypothetical revision to her opinion before deciding whether to actually make that revision. She just has to do a quick check to see if the evidence looks especially surprising at first glance.

I call this "the doorstep model". The rule-of-thumb is like a doorstep that keeps the doxastic door closed, or almost closed on the credence-first view. Once our subject has nearly closed the door on a question, so that it's not worth pushing the door closed any farther, she puts a doorstep behind it to keep it in place. Only an unusually strong nudge will dislodge the doorstep and cause her to go back to adjusting the door's position.

But the problem with the doorstep model is the doorstep itself. It's a new entity, separate from the door and its position.

If we posit a policy that gets put in place when deliberation is terminated, we admit that cognitive closure is something separate from high credence. There is a state the agent puts herself in as a result of her decision to stop deliberating, a state of obeying a rule-of-thumb for ignoring further evidence. That state is causally produced by her high credence, not constituted by it. It is an effect of her decision to terminate deliberation, and thus, indirectly, an effect of the credences that figured into it. But it itself is something new.

This "doorstep" state could still differ from a full belief. It has full belief's settled profile, but not necessarily its premise-supplying profile. Settling into a state of resistance to considerations against P needn't dispose one to rely on P as an assumption in practical reasoning. Can the credence-first philosopher still insist, then, that only the agent's high credence in P figures into her reasoning, and never P itself as an unqualified assumption?

She can, but not very contentedly. First, our credence-first philosopher has already abandoned the simple, unified Bayesian model. Instead she has embraced the kind of "bounded rationality" and slap-dash, piecemeal reliance on rules-of-thumb that we saw in §2 leads many psychologists to posit models like the adaptive toolbox. Once we've come that far, there may be little point in resisting the dualist interpretation of the literatures on cognitive closure and judgment and decision. Second, resisting counter- P considerations often looks a lot like assuming P . When we resist hypotheses and evidence that challenge what we accept, we often explain them away or dismiss them on the grounds that they must be wrong. Determined credence-first theorists can insist that this resistance doesn't assume P but instead works in some

credence-first way. But so resisting adds to their burden, which they may already want to give up carrying as just noted.

There are further challenges for a credence-first account of cognitive closure. The phenomenology of closure doesn't accompany just any old decision to stop deliberating when it's no longer worth it. Even when a question remains highly uncertain, deliberation can become too costly, because another matter becomes more pressing or because of a distraction. Yet we don't experience the same phenomenology or affect in these cases. There is no sense of having established a firm fact, and the feeling of urgency to settle the question may remain. We may even feel pulled back to those deliberations, and have to force ourselves away from them, feeling despondent and defeated rather than resolved and relieved.

And then there are cases of high credences that aren't accompanied by the "firm fact" phenomenology. As Nagel (2011) observes about the Harman-Vogel paradox, it seems a firm fact that your car is parked where you left it until a friend asks how you know it hasn't been stolen and moved. After thinking about it for a moment, you might feel very confident your car hasn't been stolen. Car-theft is extremely rare in your neighbourhood, let's suppose. But the proposition that your car is where you left it won't have the same, settled character it had initially. Is expected utility maximized by continuing to deliberate about your car's location?

4. MEMORY & METACOGNITION

Much of what we believe is stored in memory, perhaps most of it. And storage is a special concern for dualists. The Leviathan is an ungainly beast, with its two separate data structures for storing beliefs and the cognitive burden of keeping them synchronized. So it's natural to ask how the human mind ultimately stores information. Is it fundamentally graded, like Rudy the Robot? Or is it instead categorical? Or somehow both?

4.1. Empirical Background. At first glance it looks like human memory stores categorical information, sans degrees of belief. Decades of memory research suggest that confidence levels are not stored, but instead constructed at the time a memory is recalled.

Recall the availability heuristic from §2.1. The more instances of heart attacks come to mind, and the more easily they come to mind, the more probable a heart attack will seem (Tversky & Kahneman, 1973, 1974). Building on this idea, Jacoby and others proposed that *fluency*—the ease with which something is recalled or processed—affects feelings of familiarity (Jacoby & Dallas, 1981; Jacoby et al., 1989). In one particularly striking experiment, Whittlesea et al. (1990) showed subjects

seven words in rapid succession, followed by an eighth word whose text was subtly masked to make it harder to read. Subjects were asked to say “old” or “new”—whether the eighth word had been one of the previous seven or not. Subjects were more likely to say “old” when the masking was light. Evidently, the more fluently they could process the text, the more familiar it felt. (See also Jacoby & Whitehouse, 1989; Whittlesea, 1993.)

Koriat (1993, 1998) further suggested that the feeling of knowing—the sense that one knows the answer to a question even though one can’t recall it—is similarly affected by features of the recall process. Koriat identifies two features in particular: quantity and fluency. Quantity is the amount of related information that comes to mind, fluency is how easily it comes to mind. For example, Koriat (1993) showed subjects a nonsense string of letters and then asked them to recall as many of the letters as they could. If they couldn’t recall all the letters, they were asked to pick the complete string out of a list. The more letters the subjects could recall on their own (quantity), the more likely they were to think they would succeed at picking the complete string out of the list (feeling of knowing)(see also Schwartz & Smith 1997). They were also more likely to think they could correctly pick it out when the letters they had recalled on their own were recalled more quickly (fluency)—even holding the number of recalled letters constant.

Kelley & Lindsay (1993) and Koriat et al. (2006, 2008) extend this idea to subjective confidence. When subjects succeed in recalling an answer to a question, their confidence in their answer is influenced by the ease with which they recall it (fluency), and the amount of related information that comes to mind (quantity). For example, Kelley & Lindsay found that exposing subjects to correct and relevant-but-incorrect information immediately prior to a general knowledge test increased both the speed of their answers (a common proxy for fluency) and their confidence in those answers. Koriat et al. (2006) found that confidence in an answer decreased the longer subjects worked on a problem. And Koriat et al. (2008) found related results for quantity.

Summarizing these and similar results, van Overschelde (2008) writes: “Research has shown that virtually any method of making information easier to process can lead to more positive judgments.”

The punchline for our purposes is this: confidence in memory-based beliefs appears to be constructed at the time of recall, rather than stored. If you’re asked what the capital of Iceland is, the more easily the answer (Reykjavik) comes to mind, and the more related information comes to mind, the more certain you will be that your answer is correct. So your confidence that Reykjavik is the capital of Iceland doesn’t appear to be stored in memory, at least not directly. (It may be partly encoded

indirectly, in facts about your memory-state that determine how fluently you will recall the answer, or what quantity of related information will be recalled alongside.)

It would be too strong to say that confidence levels are *wholly* determined by how the recall process goes. They must also be affected by what's stored in memory. If nothing about Iceland is stored in memory, nothing will come to mind at the time of recall, and you will have virtually no confidence that the capital of Iceland is a place called 'Reykjavik'. So memory's contents play a big part in determining confidence levels too.²⁷

4.2. Philosophical Application 1: Shared Storage. One dualist-friendly moral to be drawn from these results is a solution of the parsimony problem. The Leviathan, we now see, is a grotesque caricature of dualism. Humans do not store information in two, separate belief structures, one for partial belief and the other for full belief. Full and partial belief share one, common store: long term memory.²⁸ That single, shared store is used to generate both categorical judgments and levels of confidence. In particular, degrees of belief needn't be stored. Occurrent degrees of belief can be constructed on the fly, based on features of the retrieval process like fluency and quantity. And dispositional degrees of belief are only encoded in memory indirectly, insofar as one's memory is in a state that disposes one to retrieve information with more or less fluency and quantity.

This solves several parsimony problems at once. The Leviathan faced the onerous burden of dual storage, which we now see can be eliminated. It also had the expensive cognitive task of keeping the two data structures in sync, to ensure that it generally believed *P* just when it was confident that *P*. But humans don't have this second problem either, because we don't have two separate stores to synchronize.

Even better, we can now explain why we are generally confident in *P* just when we believe *P*: because memory states conducive to judging that *P* are also conducive to doing so quickly and easily. For example, if I've studied Iceland's geography and politics in preparation for a recent exam, my memories relevant to Iceland's capital will be in a straightforward, pristine state. When asked about Iceland's capital, Reykjavik will come to mind quickly and easily and with a wealth of associated information, and I will judge that it is the capital. If I'm then prompted to assess my confidence in that answer, the fluency and quantity associated with the recall process will make me quite confident.

²⁷See, for example, Koriat et al. (2008); Koriat (2012) on the interplay between *experienced based* cues like fluency and quantity and *information based* cues, i.e. the contents of memory.

²⁸In fact there may be several different long-term memory stores, e.g. for episodic memory and for semantic memory (Nyberg & Cabeza, 2000). But this kind of division doesn't affect the point here.

An anonymous reviewer notes that this is a contingent explanation. It illuminates how we are constituted, so that high credence and full belief will typically go together. But we might have been constituted otherwise. Whereas the connection between high confidence and full belief is necessary, one might well think.

Grant for the sake of argument that the connection between full belief and high confidence is necessary. What is being explained here is not that necessary connection. That would be explained by the functional roles associated with ‘belief’ and ‘confidence’. Rather, what is being explained is how human cognition manages to instantiate those roles, which posit that connection and make it necessary. The special challenge for the dualist is to offer a plausibly economical story to match the credence-first monist’s. She has to explain how humans achieve synchronization between the states that, in us, realize full and partial belief, without positing absurdly demanding or elaborate cognitive machinery. And this she has done.

4.3. Philosophical Application 2: A Quasi-Categorical-First Hypothesis. An ambitious but naive categorical-first philosopher might try to go further, mounting the following argument. The fact that confidence in memory-based judgment is assessed at the time of recall turns the credence-first picture on its head. Far from being ontologically prior, partial belief is actually dependent on full belief, at least in these memory-based cases. These partial beliefs are generated based on how the process of recalling categorical information about Iceland goes. Thus full beliefs comes first! Partial beliefs only emerge as a sort of byproduct of the storage and retrieval process for full beliefs.

Even if this argument succeeded, it would not show that partial beliefs supervene on full beliefs. In fact, it would show that they don’t. On the model we are entertaining, there is one process for retrieving contents from memory and basing categorical judgments on them, and a second process which monitors the performance of the first to generate a confidence level. The second process needs the first to function, and maybe this dependence is so strong that we should say these confidence levels are partly constituted by the categorical judgment process. But they would still only be *partly* constituted by that process. They would also be partly constituted by the second process, the one that monitors the first. It would be wrong to say our partial beliefs are nothing over and above our full beliefs.

Nevertheless, it would be interesting to learn that our full beliefs “precede” our degrees of belief in this weaker sense. At the very least, it would starkly illustrate one way in which credence-first philosophers were wrong to put credences first. So let’s explore this *quasi*-categorical-first argument further.

The argument is much too naive. A respectable argument in this vein can be made, but it is significantly more tentative. For there is a quasi-*credence*-first interpretation of the empirical findings. And to rule it out, we have to rely on more speculative considerations.

We've seen evidence that confidence levels are not directly stored, but constructed at the time of recall. The same may be true of categorical judgments, however.

Suppose I learn that it's actually a common misconception that Reykjavik is the capital of Iceland. Keflavik was actually made Iceland's capital by a little-known act of parliament in 1978. Now I have two relevant pieces of information stored in memory. For at least a little while, I will continue to store the information that Iceland is the capital; I've thought so for so long that it will take some time to scrub that item from memory. But in the meantime, I also store the information that this is a common misconception, thanks to a largely overlooked act of parliament. If asked what the capital of Iceland is, both items will come to mind, and I will judge that Keflavik is the capital, not Reykjavik.

Thus stored information does not a full belief make. I can store the information that Reykjavik is the capital without believing it. Whether I believe Reykjavik is the capital depends on what I am disposed to make of that stored information upon retrieving it from memory.

Here the credence-first philosopher has a wedge she might drive further. How do we decide what to make of the information we retrieve from memory? Especially when there's a clash between retrieved items (Reykjavik vs. Keflavik), there's a puzzle about how things are resolved to reach a final answer. And one natural view is that we rely on partial beliefs to resolve things. We first assess the plausibility of various explanations for having the clashing items in memory, and only then do we reach a categorical conclusion. For example, since my memory about Reykjavik being the capital is probably an instance of the common misconception, I will conclude that Keflavik is the capital, not Reykjavik.

In general, the credence-first philosopher sees the steps of memory-based judgment going something like this:

- (1) Search the contents of memory for relevant information.
- (2) Form confidence levels in the available answers based on:
 - (a) what information is retrieved, and
 - (b) features of the retrieval process like fluency and quantity.
- (3) Based on those confidence levels, decide what categorical judgment to make.

On this view full and partial belief are *both* constructed at the time of recall. More exactly, *occurrent* partial beliefs (confidence levels) and *occurrent* full beliefs (categorical judgments) are constructed. But, importantly, the *occurrent* full beliefs are constructed on the basis of the *occurrent* partial beliefs, as a sort of cognitive decision.²⁹ So partial belief actually precedes full belief, in a sense.

We have two competing pictures here. The first supports the quasi-categorical-first view. It posits a mechanism for forming categorical judgments on the basis of retrieved information that does not rely on levels of confidence. Confidence levels are instead formed based on the performance of the retrieval-and-categorical-judgment process. The second picture supports a quasi-credence-first view. It says the mechanism that forms categorical judgments on the basis of retrieved information does require that confidence levels be assessed first. These confidence levels are always used as the basis for categorical conclusions drawn from stored information.

Both pictures are compatible with the finding we began with: that confidence in a memory-based judgment is affected by features of the recall process like fluency and quantity. They just differ on when and how that effect arises. The first picture says the effect on confidence happens after a categorical judgment is rendered. The second picture says instead that the effect on confidence happens before any categorical judgment is rendered.

There are some broad, tentative reasons to suspect the first picture is correct, at least in some cases. It posits a simpler, more optimal design in some ways. The bulk of the quotidian beliefs that guide us every day come from percepts and memories we have no reason to doubt, and no need to form confidence levels in. The information I rely on to execute my morning routine—to make breakfast and check my email—doesn't need accompanying confidence levels. So always assessing those levels would be a waste of cognitive resources. A setup that only assessed confidence levels when necessary would be more efficient. It could treat the contents of memory and perception as true by default, judging that *P* when *P* is delivered by memory or perception, provided the delivery happens smoothly (and the stakes are not too high, etc.). When the delivery isn't smooth—when recall is disfluent or delivers conflicting information—then more expensive, confidence-involving processing could kick in.

There are empirical grounds for this view. First, disfluency is known to function as a sort of alarm bell that triggers more effortful and time-consuming processing, so-called “System 2” type processing (Alter et al., 2007). Second, Stanovich (2009) argues extensively that humans are “cognitive misers”. We generally avoid this kind of effortful and time-consuming processing unless we have cause to engage in it. Third,

²⁹See Kaplan (1983); Koriat (2012); Levi (1967, 1980); Maher (1993) for related views.

on the received view the kind of self-monitoring that tracks features like fluency involves the prefrontal cortex (Shimamura, 2008), and is also probably fairly new, evolutionarily (Metcalf, 2008). So other animals appear to have a way of using and reconciling retrieved information without relying on confidence levels formed based on fluency and quantity. Humans presumably possess the same means, but also possess the capacity to monitor its fluency and quantity, since these can be useful when more judicious, more expensive processing is called for.

These considerations provide some support, but they are hardly conclusive—really only tentative at best. We simply don't know enough to make a strong case for one interpretation over the other.

Still, it should be clear that a certain reductivist tendency in credence-first epistemology is too hasty. It's very much an open empirical question what mechanisms undergird human memory, and memory-based judgment. On top of that come the difficult conceptual questions about how to apply the notions of full and partial belief to those mechanisms. Bayesian epistemologists might be tempted to declare that credence comes first, and that Bayesian epistemology and decision theory complete our normative theorizing. But that stance is badly premature. The small taste we've gotten of the difficulties in this area should be enough to demonstrate that much.

5. STEPPING BACK

We started by noting the different functional roles typically associated with full and partial belief. Full beliefs provide premises that are treated as givens in reasoning, while partial beliefs are used as weights. Full beliefs also have a distinctive, settled character, being resistant to disruption. We proceeded to explore three literatures in cognitive psychology to find where and how these roles might be fulfilled.

Psychologists studying judgment and decision have largely converged on the view that we don't rely exclusively on the kind of weighting associated with partial beliefs. It's not just that we don't always meet classical ideals like Bayes' theorem and expected utility maximization. It's that the methods we use instead rely on the kinds of givens full beliefs are supposed to provide. This pattern has held for over four decades now, from the first work in the 1970s on heuristics like representativeness and elimination-by-aspects, through the 1980s and 1990s when the "adaptive toolbox" paradigm dominated, and into present work on the evidence accumulation model. These models don't rely exclusively on categorical inputs, however. Many of them combine categorical premises with probabilistic weights. Taken at face value then, they support a dualist ontology. Credence-first philosophers may resist taking them

at face value, but we saw that this resistance comes with costs. It adds computational labour to a model whose whole point is to save on such labour.

What about full belief's second distinguishing characteristic, its settled character?

Psychologists have long posited settled doxastic states, which we seek to maintain and feel uncomfortable when disturbed from. The most carefully studied and well-confirmed such posit is the state of cognitive closure, the state where a belief has crystallized into a firm fact from the subject's point of view. The theory of cognitive closure makes numerous confirmed predictions, of which we surveyed just a few. But most notably, it predicts a kind of "closed mindedness", where subjects become resistant to further deliberation and new evidence.

The theory of cognitive closure also predicts that subjects are more motivated to achieve such closure when a decision is impending. This connects cognitive closure to the research on judgment and decision, suggesting that there is a state which plays both roles classically associated with full belief: a state where the subject becomes resistant to considerations counter to P , and where she will use P as an assumption in reasoning.

What about dualism's challenges? Recall that dualists face a difficult dilemma. If full and partial belief agree perfectly in the governance of inference, action, and assertion, then one of them is extraneous. If instead they differ, then we get conflicting predictions. And some of these conflicts are quite troublesome. Acting on your full belief that Reykjavik is the capital of Iceland could lead you to wager your life on it in exchange for a penny.

Resources for resolving this dilemma are also to be found in the work on judgment and decision and on cognitive closure. Full and partial belief would recommend conflicting courses of action if both modes of belief were active at the same time for the same proposition. But we saw in the judgment and decision literature that P is always treated as either a given or as having a certain weight, not both. Some mechanisms, like evidence accumulation, can treat P as given and Q as weighted, but only when $P \neq Q$. So the dualist resolution begins with this observation: we do not have an *occurrent* full belief in P and an *occurrent* partial belief about P at the same time. We do simultaneously have full and partial beliefs about one and the same propositions *dispositionally*. But only one of these dispositional states will be active at any given moment.

So what keeps you from betting your life on Reykjavik being the capital of Iceland? This turns on the question when each kind of dispositional belief becomes occurrent. We didn't develop a full, detailed answer to this general question. But we did note that extreme stakes can trigger more deliberate and careful reasoning, where a believed

proposition is no longer taken as a given, but instead treated as highly probable. We also noted that numerous other factors surveyed by Nagel (2008, 2010, 2011) can trigger such shifts, which Nagel applies to other instances of the dualist dilemma like the Harman-Vogel paradox. Exactly how they might apply to further instances, like the preface paradox, was left as an open question for future research.

A second challenge for dualism is parsimony. Not only does dualism posit two states where credence-first philosophers ultimately have only one. It also seems to posit extra storage. Memory must somehow record what the agent fully believes in addition to what she partially believes. Such dual storage looks especially suspicious given the tight coordination full and partial belief seem to enjoy. Generally speaking, people believe something when and only when they are confident of it. Isn't the simplest explanation that full belief just is strong partial belief?

The literature on memory and confidence resolves the puzzle. Neither full nor partial beliefs are stored explicitly. The image of the Leviathan, with its two, separate belief-boxes is misleading. Rather, memory stores information which gets turned into categorical judgments and confidence levels on the fly, as needed. When the question whether *P* arises, relevant information is retrieved and a judgment about *P* is formed on that basis. A confidence-level in *P* can also be rendered, based on how smooth and rich the retrieval process is. The reason full belief in *P* and strong partial belief in *P* go together is that states of memory that tend to yield judgments that *P* also tend to do so smoothly and richly, so that confidence in that judgment comes along with it. There is no special task of keeping full and partial belief in sync. There is only the task of managing the information in one, shared memory store.

We also saw that there may even be an argument here that full belief is prior to partial belief in an interesting sense. The mechanisms that generate confidence levels in memory-based judgments seem to be parasitic on the mechanisms that generate categorical judgments. So confidence may be a byproduct of categorical judgment, at least for memory-based beliefs. But we also saw that there is an alternative interpretation, where things are reversed. It's possible that subjects first form confidence levels based on the smoothness and richness of the retrieval process, and then generate categorical judgments based on those confidence levels. We saw some reasons in favour of the first interpretation, but they could only be tentative given the current state of research.

Stepping back, we see that the dualist ontology of belief enjoys support of various kinds from various empirical sources. Tentatively at least, we ought to conclude that belief in psychology's ontology is dual.

REFERENCES

- Alter, Adam L., Daniel M. Oppenheimer, Nicholas Epley & Rebecca N. Eyre. 2007. Overcoming Intuition: Metacognitive Difficulty Activates Analytic Reasoning. *Journal of Experimental Psychology: General* 136(4): 569–576.
- Bettman, James R., Eric J. Johnson & John W. Payne. 1990. A Componential Analysis of Cognitive Effort in Choice. *Organizational Behavior and Human Decision Processes* 45(1): 111–139.
- Brown, Jessica. 2011. Fallibilism and the Knowledge Norm for Assertion and Practical Reasoning. In Jessica Brown & Herman Cappelen (eds.), *Assertion: New philosophical essays*, Oxford University Press.
- Brown, Jessica. 2012. Assertion and Practical Reasoning: Common or Divergent Epistemic Standards? *Philosophy and Phenomenological Research* 84(1): 123–157.
- Carnap, Rudolf. 1971. Inductive Logic and Rational Decisions. In Rudolf Carnap & Richard C. Jeffrey (eds.), *Studies in inductive logic and probability*, vol. 1, 5–31. University of California Press.
- Christensen, David. 2004. *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford.
- Coombs, C. H., H. Raiffa & R. M. Thrall. 1954. Some Views on Mathematical Models and Measurement Theory. *Psychological Review* 61(2): 132–144.
- Cresto, Eleonora. 2010. On Reasons and Epistemic Rationality. *The Journal of Philosophy* 107(6): 326–330.
- Davies, Mark & Dee Gardner. 2010. Frequency Dictionary of American English .
- Dodd, Dylan. ????. Belief and Certainty.
- Douven, Igor. 2006. Assertion, Knowledge, and Rational Credibility. *The Philosophical Review* 115(4): 449–485.
- Douven, Igor. 2008. Knowledge and Practical Reasoning. *Dialectica* 62(1): 101–118.
- Douven, Igor. 2009. Assertion, Moore, and Bayes. *Philosophical Studies* 144(3): 361–375.
- Easwaran, Kenny. 2015. Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probabilities. *Noûs* .
- Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Stanford University Press.
- Foley, Richard. 1993. *Working Without a Net: A Study of Egocentric Epistemology*. Oxford University Press.
- Foley, Richard. 2009. Beliefs, Degrees of Belief, and the Lockean Thesis. In Franz Huber & Christoph Schmidt-Petri (eds.), *Degrees of belief*, 37–47. Springer.
- Frankish, Keith. 2009. Partial Belief and Flat-Out Belief. In Franz Huber & Christoph Schmidt-Petri (eds.), *Degrees of belief*, Dordrecht: Springer Netherlands.

- Fumerton, Richard. 1995. *Metaepistemology and Skepticism*. Rowman & Littlefield.
- Gigerenzer, Gerd & Daniel G. Goldstein. 1996. Reasoning the Fast and Frugal Way: Models of Bounded Rationality 103(4): 650–669.
- Gigerenzer, Gerd, Peter M. Todd & ABC Research Group. 1999. *Simple Heuristics that Make Us Smart*. Oxford University Press.
- Goldstein, Daniel G. & Gerd Gigerenzer. 2002. Models of Ecological Rationality: The Recognition Heuristic. *Psychological Review* 109(1): 75–90.
- Good, I. J. 1967. On the Principle of Total Evidence. *Philosophy of Science* 17(4): 319–321.
- Hacking, Ian. 1967. Slightly More Realistic Personal Probability. *Philosophy of Science* 34(4): 311–325.
- Harman, Gilbert. 1973. *Thought*. Princeton University Press.
- Hawthorne, James. 2009. The Lockean Thesis and the Logic of Belief. In Franz Huber & Christoph Schmidt-Petri (eds.), *Degrees of belief*, 49–74. Springer: Springer.
- Hawthorne, James & Luc Bovens. 1999. The Preface, the Lottery, and the Logic of Belief. *Mind* 108(430): 241–264.
- Hawthorne, John & Jason Stanley. 2008. Knowledge and Action. *The Journal of Philosophy* 105(10): 571–590.
- Heider, Fritz. 1958. *The Psychology of Interpersonal Relations*. Psychology Press.
- Huber, Oswald. 1980. The influence of some task variables on cognitive operations in an information-processing decision model. *Acta Psychologica* 45(1-3): 187–196.
- Jacoby, Larry L. & Mark Dallas. 1981. On the Relationship Between Autobiographical Memory and Perceptual Learning. *Journal of Experimental Psychology: General* 110(3): 306–340.
- Jacoby, Larry L., Colleen M. Kelley & Jene Dywan. 1989. Memory Attributions. In Henry L. Roedinger III & Fergus I. M. Craik (eds.), *Varieties of memory and consciousness: Essays in honour of endel tulving*, 391–422. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Jacoby, Larry L. & Kevin Whitehouse. 1989. An Illusion of Memory: False Recognition Influenced by Unconscious Perception. *Journal of Experimental Psychology: General* 118(2): 126–135.
- Jeffrey, Richard C. 1970. Dracula Meets Wolfman: Acceptance vs. Partial Belief. In *Induction, acceptance and rational belief*, 157–185. Dordrecht: Springer Netherlands.
- Jeffrey, Richard C. 1983. *The Logic of Decision*. Chicago, IL: University of Chicago Press: The University of Chicago Press 2nd edn.
- Kagan, Jerome. 1972. Motives and development. *Journal of Personality and Social Psychology* 22(1): 51–66.

- Kahneman, Daniel & Amos Tversky. 1973. On the Psychology of Prediction. *Psychological Review* 80(4): 237–251.
- Kahneman, Daniel & Amos Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2): 263.
- Kaplan, Mark. 1983. Decision Theory as Philosophy. *Philosophy of Science* 50(4): 549–577.
- Kaplan, Mark. 1996. *Decision Theory as Philosophy*. Cambridge University Press.
- Kelley, Colleen M. & D. Stephen Lindsay. 1993. Remembering Mistaken for Knowing: Ease of Retrieval as a Basis for Confidence in Answers to General Knowledge Questions. *Journal of Memory and Language* 32(1): 1–24.
- Koriat, Asher. 1993. How Do We Know That We Know? The Accessibility Model of the Feeling of Knowing. *Psychological Review* 100(4): 609–639.
- Koriat, Asher. 1998. Illusions of Knowing: The Link Between Knowledge and Metaknowledge. In Vincent Y. Yzerbyt, Guy Lories & Benoit Dardenne (eds.), *Metacognition: Cognitive and social dimensions*, 16–34.
- Koriat, Asher. 2012. The Self-consistency Model of Subjective Confidence. *Psychological Review* 119(1): 80–113.
- Koriat, Asher, Hilit Ma'ayan & Ravit Nussinson. 2006. The Intricate Relationships Between Monitoring and Control in Metacognition: Lessons for the Cause-and-effect Relation Between Subjective Experience and Behavior. *Journal of Experimental Psychology: General* 135(1): 36–69.
- Koriat, Asher, Ravit Nussinson, Herbert Bless & Nira Shaked. 2008. Information-Based and Experience-Based Metacognitive Judgments: Evidence from Subjective Confidence. In John Dunlosky & Robert A. Bjork (eds.), *Handbook of metamemory and memory*, 117–135. New York, NY: Taylor & Francis.
- Kruglanski, Arie W. 1989. *Lay Epistemics and Human Knowledge* Cognitive and Motivational Bases. Springer Science & Business Media.
- Kruglanski, Arie W. 2004. *The Psychology of Closed Mindedness*. New York, NY: Psychology Press.
- Kruglanski, Arie W. & Woo Young Chun. 2008. Motivated Closed-Mindedness and Its Social Consequences . In James Y. Shah & Wendi L. Gardner (eds.), *Handbook of motivation science*, 84–99.
- Kruglanski, Arie W. & Donna M. Webster. 1991. Group Members' Reactions to Opinion Deviates and Conformists at Varying Degrees of Proximity to Decision Deadline and of Environmental Noise. *Journal of Personality and Social Psychology* 61(2): 212–225.

- Kruglanski, Arie W. & Donna M. Webster. 1996. Motivated Closing of the Mind: "Seizing" and "Freezing". *Psychological Review* 103(2): 263–283.
- Kruglanski, Arie W., Donna M. Webster & Adena Klem. 1993. Motivated Resistance and Openness to Persuasion in the Presence or Absence of Prior Information. *Journal of Personality and Social Psychology* 65(5): 861–876.
- Kyburg, Henry E. 1961. *Probability and the Logic of Rational Belief*. Middletown, CT: Wesleyan University Press.
- Lance, Mark Norris. 1995. Subjective probability and acceptance. *Philosophical Studies* 77(1): 147–179.
- Lee, Michael D. & Tarrant D R Cummins. 2004. Evidence Accumulation in Decision Making: Unifying The "Take the Best" and The "Rational" Models. *Psychonomic Bulletin & Review* 11(2): 343–352.
- Leitgeb, Hannes. 2013. Reducing Belief Simpliciter to Degrees of Belief. *Annals of Pure and Applied Logic* 164(12): 1338–1389.
- Leitgeb, Hannes. 2014. The Stability Theory of Belief. *The Philosophical Review* 123(2): 131–171.
- Lerner, Jennifer S. & Philip E. Tetlock. 1999. Accounting for the Effects of Accountability. *Psychological Bulletin* 125(2): 255–275.
- Levi, Isaac. 1967. *Gambling With Truth*. New York, NY: Alfred A. Knopf.
- Levi, Isaac. 1980. *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*. Cambridge, MA: The MIT Press.
- Lewis, David. 1970. How to Define Theoretical Terms. *The Journal of Philosophy* 67(13): 427–446.
- Lewis, David. 2010. Why conditionalize? In *Papers in metaphysics and epistemology*, 403–407. Cambridge: Cambridge University Press.
- Lin, Hanti & Kevin T. Kelly. 2012. Propositional Reasoning that Tracks Probabilistic Reasoning. *Journal of Philosophical Logic* 41(6): 957–981.
- Luce, Robert Duncan & Howard Raiffa. 1957. *Games and Decisions*. John Wiley and Sons, Inc.
- Maher, Patrick. 1993. *Betting on Theories*. New York, NY: Cambridge University Press.
- Makinson, David C. 1965. The paradox of the preface. *Analysis* 25(6): 205–207.
- Metcalf, Janet. 2008. Evolution of Metacognition. In John Dunlosky & Robert A. Bjork (eds.), *Handbook of metamemory and memory*, 29–45. New York, NY: Psychology Press.
- Nagel, Jennifer. 2008. Knowledge Ascriptions and the Psychological Consequences of Changing Stakes. *Australasian Journal of Philosophy* 86(2): 279–294.

- Nagel, Jennifer. 2010. Epistemic Anxiety and Adaptive Invariantism. *Philosophical Perspectives* 24: 407–435.
- Nagel, Jennifer. 2011. The Psychological Basis of the Harman-Vogel Paradox. *Philosophers' Imprint* 11(5): 1–28.
- Newell, Allen & Herbert A. Simon. 1979. Human problem solving. Prentice-Hall.
- Newell, Ben R., Patrick Collins & Michael D. Lee. 2007. Adjusting the Spanner: Testing an Evidence Accumulation Model of Decision Making. *proceedings of the 29th Annual Conference of the Cognitive Science Society* 533–538.
- Newell, Ben R. & Michael D. Lee. 2011. The Right Tool for the Job? Comparing an Evidence Accumulation and a Naive Strategy Selection Model of Decision Making. *Journal of Behavioral Decision Making* 24(5): 456–481.
- Nyberg, Lars & Roberto Cabeza. 2000. Brain Imaging of Memory. In Endel Tulving & Fergus I. M. Craik (eds.), *The oxford handbook of memory*, 501–519.
- Oaksford, Mike & Nick Chater. 1998. *Rationality in an Uncertain World*. New York, NY: Taylor & Francis.
- Oaksford, Mike & Nick Chater. 2007. *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford University Press.
- Oaksford, Mike & Nick Chater. 2009. Précis of Bayesian Rationality: The Probabilistic Approach to Human Reasoning. *Behavioral and Brain Sciences* 32(1): 69–120.
- Payne, John W. & James R. Bettman. 2004. Walking with the Scarecrow: The Information-processing Approach to Decision Research. In Derek J. Koehler & Nigel Harvey (eds.), *Blackwell handbook of judgment and decision making*, 110–131. Malden, MA: Blackwell.
- Payne, John W., James R. Bettman & Eric J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge University Press.
- Pollock, John L. 2006. *Thinking about Acting: Logical Foundations for Rational Decision Making*. Oxford University Press, USA.
- Ross, Jacob & Mark Schroeder. 2014. Belief, Credence, and Pragmatic Encroachment. *Philosophy and Phenomenological Research* 88(2): 259–288.
- Savage, Leonard J. 1954. *The Foundations of Statistics*. Hoboken, NJ: Wiley.
- Schiffer, Stephen. 2007. Interest-Relative Invariantism. *Philosophy and Phenomenological Research* 75(1): 188–196.
- Schwartz, Bennett L. & Steven M. Smith. 1997. The Retrieval of Related Information Influences Tip-of-the-Tongue States. *Journal of Memory and Language* 36(1): 68–86.
- Schwitzgebel, Eric. 2015. Belief. <http://plato.stanford.edu/entries/belief/>.

- Shimamura, Arthur P. 2008. A Neurocognitive Approach to Metacognitive Monitoring and Control. In John Dunlosky & Robert A. Bjork (eds.), *Handbook of metamemory and memory*, 373–389. New York, NY: Psychology Press.
- Shoham, Yoav. 1987. A Semantical Approach to Nonmonotonic Logics. In Matthew L. Ginsberg (ed.), *Readings in Nonmonotonic Reasoning*, 227–250. Los Altos, CA: Morgan Kaufman.
- Stalnaker, Robert C. 1984. *Inquiry*. Cambridge, MA: MIT Press: The MIT Press.
- Stanovich, Keith. 2011. *Rationality and the Reflective Mind*. Oxford University Press.
- Stanovich, Keith E. 2009. *What Intelligence Tests Miss*. New Haven, CT: Yale University Press.
- Sturgeon, Scott. 2008. Reason and the Grain of Belief. *Noûs* 42(1): 139–165.
- Tversky, Amos. 1969. Intransitivity of Preferences. *Psychological Review* 76(1): 31–48.
- Tversky, Amos. 1972. Elimination by Aspects: A Theory of Choice. *Psychological Review* 79(4): 281–299.
- Tversky, Amos & Daniel Kahneman. 1973. Availability: a Heuristic for Judging Frequency and Probability. *Cognitive Psychology* 5(2): 207–232.
- Tversky, Amos & Daniel Kahneman. 1974. Judgment Under Uncertainty: Heuristics and Biases. *Science* 185(4157): 1124–1131.
- Tversky, Amos & Daniel Kahneman. 1983. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90(4): 293–315.
- Van Overschelde, J P. 2008. Metacognition: Knowing About Knowing. In John Dunlosky & Robert A. Bjork (eds.), *Handbook of metamemory and memory*, New York, NY: Psychology Press.
- Vogel, Jonathan. 1990. Are there Counterexamples to the Closure Principle? In Michael D. Roth & Glenn Ross (eds.), *Doubting: Contemporary perspectives on skepticism*, 13–27. Springer: Springer Netherlands.
- Vogel, Jonathan. 2000. Reliabilism Leveled. *The Journal of Philosophy* 97(11): 602–623.
- Wason, Peter C. 1966. Reasoning. In Brian M Foss (ed.), *New horizons in psychology*, 135–151. Penguin.
- Webster, Donna M. 1993. Motivated Augmentation and Reduction of the Overattribution Bias. *Journal of Personality and Social Psychology* 65(2): 261–271.
- Webster, Donna M. & Arie W. Kruglanski. 1994. Individual Differences in Need for Cognitive Closure. *Journal of Personality and Social Psychology* 67(6): 1049–1062.
- Webster, Donna M., Linda Richter & Arie W. Kruglanski. 1996. On Leaping to Conclusions When Feeling Tired: Mental Fatigue Effects on Impressional Primacy. *Journal of Experimental Social Psychology* 32(2): 181–195.

- Weisberg, Jonathan. 2010. Bootstrapping in General. *Philosophy and Phenomenological Research* 81(3): 525–548.
- Weisberg, Jonathan. 2012. The Bootstrapping Problem. *Philosophy Compass* 7(9): 597–610.
- Weisberg, Jonathan. 2013. Knowledge in Action. *Philosophers' Imprint* 13(22): 1–23.
- Whittlesea, Bruce W. A. 1993. Illusions of Familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19(6): 1235–1253.
- Whittlesea, Bruce W. A., Larry L. Jacoby & Krista Girard. 1990. Illusions of Immediate Memory: Evidence of an Attributional Basis for Feelings of Familiarity and Perceptual Quality. *Journal of Memory and Language* 29(6): 716–732.
- Wu, George, Jiao Zhang & Richard Gonzalez. 2004. Decision Under Risk. In Derek J. Koehler & Nigel Harvey (eds.), *Blackwell handbook of judgment and decision making*, 399–423. Blackwell Publishing.