Belief and Evidence

by

Ginger Schultheis

B.A., Reed College (2011)

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

Signature redacted

Signature of author: ______________

Department of Linguistics and Philosophy
August 31, 2018

Signature redacted

Certified by: ______________

Roger White
Professor of Philosophy
Thesis Supervisor

Signature redacted

Accepted by: ______________

Bradford Skow
Laurance S. Rockefeller Professor of Philosophy
Chair of Committee on Graduate Students

Belief and Evidence

by

Ginger Schultheis

Submitted to the Department of Linguistics and Philosophy on August 31st 2018
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Philosophy

Abstract

Chapter 1, 'Living on the Edge: Against Epistemic Permissivism,' argues that Epistemic Permissivists face a special problem about the relationship between our first- and higher-order attitudes. They claim that rationality often permits a *range* of doxastic responses to the evidence. Given plausible assumptions about the relationship between your first- and higher-order attitudes, you can't stably be on the edge of the range, so there can't be a range at all. Permissivism, at least as it has been developed so far, can't be right. I consider some new ways of developing Permissivism, but each has problems of its own.

Chapter 2, 'Belief and Probability,' argues that rational belief doesn't reduce to subjective probability. Under the right circumstances, I argue, acquiring conflicting evidence can defeat your entitlement to believe a certain hypothesis without probabilistically disconfirming that hypothesis. I consider three probabilistic theories of rational belief—a simple threshold view, Hannes Leitgeb's stability theory, and a new theory involving imprecise credence—and show that none of them can account for the cases I describe.

Chapter 3, 'Can We Decide to Believe?', takes up the question of whether we can decide to believe. There are two main arguments for the conclusion that believing at will is impossible, which I call the *retrospective argument* and the *aim-of-belief argument*, respectively. Neither, I argue, demonstrates that believing at will is impossible in all cases. The retrospective argument leaves open the possibility of believing at will in acknowledged permissive cases; the aim-of-belief argument leaves open the possibility of believing at will when credal attitudes are imprecise.

Thesis Supervisor: Roger White
Title: Professor of Philosophy

Acknowledgements

I would like to thank Alex Byrne, David Builes, Bob Stalnaker, Vera Flocke, Miriam Schoenfield, Caspar Hare, Sophie Horowitz, Quinn White, Dan Greco, and Brendan de Kenessey. Special thanks to David Boylan, Matthew Mandelkern, and Milo Phillips-Brown for being great friends and co-authors, as well as to Bernhard Salow and Kevin Dorst for introducing me to epistemology and teaching me so much about it over the past few years. I couldn't have asked for a better committee. Thank you, Steve, Jack, Roger, and Kieran. Finally, I am eternally grateful to my parents and to my wife Caroline for their endless support over the last few years.

Contents

Introduction

My dissertation comprises three papers. The first is about epistemic permissivism; the second is about the relationship between belief and credence; and the third is about whether we can believe at will. I briefly summarize them here.

A body of evidence is *permissive* just if there may be two fully rational people who both possess that evidence yet have different opinions on some topic. Epistemic Permissivism is the view that some bodies of evidence are permissive. Permissivism is something of the default position, and rightly so: common sense tells us that disagreement among fully rational agents with the same evidence is all too common.

In the first essay of my dissertation, entitled 'Living on the Edge: Against Epistemic Permissivism,' I argue that we must part with common sense on this issue. I begin with the question of how one's first-order opinions ought to be informed and regulated by one's beliefs *about one's evidence*—in particular, one's beliefs about which attitudes the evidence permits. I defend a certain minimal connection: you shouldn't adopt an attitude that's risky by your own lights if you know of a safer option. More precisely, rational agents do not hold *dominated* attitudes: if you're sure that it's rational to be agnostic about P, and not sure that it's rational to be confident in P, then you shouldn't be confident in P. If we accept this, I argue, Epistemic Permissivism can't be right.

In the second essay, 'Belief and Probability,' I turn to the relationship between belief and subjective probability. The great hope of probabilistic theories of evidence is to show that all facts about rational belief are determined by probabilistic facts. Any feature of evidence that rational belief seems sensitive to must be captured probabilistically. My first dissertation paper challenges this thought. I argue that no existing probabilistic theory of evidence provides an adequate account of *conflicting evidence*, the sort of evidence you have when two of your sources of information contradict each other, such as when doctors offer conflicting diagnoses. Under the right circumstances, acquiring conflicting evidence can defeat your entitlement to believe a certain hypothesis without *disconfirming* that hypothesis—that is, without making it any less likely. I consider three probabilistic theories, each of which is intended to capture a different set of features of evidence. Each theory fails because it is not equipped to handle conflicting evidence.

The third essay, 'Can We Decide to Believe?', takes up the question of whether we can believe at will. Bernard Williams famously argued that believing at will is impossible—not just a contingent psychological disability, but a metaphysically necessary limitation grounded in the nature of belief—and he is the source of most contemporary discussion of this issue. I distinguish two arguments in Williams's seminal essay, which I call the *retrospective argument* and the *aim-of-belief argument*, respectively. I argue that neither arguments succeeds in demonstrating that believing at will is impossible in all cases. The retrospective argument leaves open the possibility of believing at will in acknowledged permissive cases; the aim-of-belief argument leaves open the possibility of believing at will when one's credal attitudes are imprecise.

Chapter 1

Living on the Edge: Against Epistemic Permissivism

Matt and Abby are members of a jury for a murder case. They have all the same evidence and review it separately. When they convene to discuss their conclusions, they discover that they disagree. Matt is confident that Jones is innocent; Abby is confident that Jones is guilty. When they learn of their disagreement, what do they discover about themselves? Clearly, they learn that one of them is confident in a falsehood: Either Jones is guilty or he isn't. But do they also learn that one of them has been less than fully rational, that one has failed to properly assimilate the evidence before him?

Proponents of **Epistemic Uniqueness** say yes: Given your total evidence, there is a unique rational doxastic attitude that you can take to any proposition.[1] Matt and Abby have the same evidence, but have different levels of confidence in the proposition that Jones is guilty (*Guilty*). Uniqueness entails that at most one of them is fully rational.

Proponents of **Epistemic Permissivism** say no. Matt and Abby may simply have different standards of reasoning. For example, Matt may tend to favor simple hypotheses, and Abby may tend to favor complex, more explanatory hypotheses. A permissivist thinks both standards of reasoning may be perfectly rational even

[1]This formulation is from White (2005).

though they sanction very different responses to the evidence.[2]

Permissivists paint a simple, attractive picture of rationality. There are some general, justifiable rules—e.g., Conditionalization, Probabilism, and the Principal Principle—and a wide range of permissible starting points. So long as you begin at one of these, and you follow the rules, you're doing fine. Permissivists charge that Uniquers paint a much more complicated, and metaphysically loaded, picture of rationality.[3] Uniquers must say everything Permissivists say, and much more—that there is a *unique* rational starting point, a *unique* rational credence in *Guilty*, and so forth.

I argue that Permissivists face a special challenge about the interaction between our first- and higher-order attitudes. They claim that rationality often permits a *range* of credences in a certain proposition. Yet given certain plausible assumptions about the relationship between our first- and higher-order attitudes, you cannot adopt a credence on the edge of that range. But Permissivism says that for some such range, *any* credence in that range is rational. So Permissivism is false. I consider new ways of developing Permissivism to avoid this argument, but they have problems of their own.[4] I conclude that Permissivism is not as simple as advertised, and without new motivations, it's not very attractive either.

On to the argument.

1.1 Dominance

Suppose that Permissivism is true and that Matt and Abby's evidence rationalizes any credence between, say, .3 and .7 in the proposition that Jones is guilty

[2]See Greco and Hedden (2015), Horowitz (2013), and White (2005) for arguments against versions of Epistemic Permissivism. See Ballantyne and Coffman (2012), Douven (2009), Kelly (2014), Kopec (2015), Meacham (2014), and Schoenfield (2014) for defenses. See Kopec and Titelbaum (2015) for an overview of various arguments for and against.

[3]See e.g., Schoenfield (2014) and Kelly (2014) for this criticism.

[4]Strictly speaking, someone could be a Permissivist yet deny that there is ever a *range* of permissible credences in any proposition—you might think, say, that there are just two permissible credence in the proposition that Jones is guilty, .3 and .7. But the standard motivations—that there are some general, justifiable rules, and a wide range of starting points—strongly suggest that there will be a wide range of permissible credences in most propositions. If the Permissivist denies this, she owes us a general story about why there can never be such a range. Since no story like this has been told, I ignore this complication in this paper.

(*Guilty*). Quite plausibly, Matt is not always in a position to discern the *exact* boundaries of the permissible range.[5] Compare: He isn't able to determine the exact height of a tree some distance off just by looking—his eyesight is nowhere near that good. Similarly, by reflection alone, he can't reliably determine the value of the upper and lower bounds of the permissible range to the nearest (say) .01 degrees of confidence—the evidence is too complex, and his powers of reasoning are nowhere near that good. Even if he truly believes that the lower bound is .3, he's merely guessing—for all he knows, the lower bound is .31 or .29. He can't reliably distinguish the actual case from one in which the lower bound is slightly higher or slightly lower.

Matt ought to know this about himself. He ought to know that if he were to guess the exact values of the upper and lower bound, he'd likely err. So, if the lower bound is .3, he shouldn't be *certain* that the lower bound is exactly .3. Compare: Since you know that you can't reliably guess the exact height of the tree some distance off, you shouldn't be certain that the tree is exactly 667 inches tall, even if it is—you should recognize that it might be slightly taller or slightly shorter. Similarly, Matt should acknowledge that the lower bound might well be slightly greater than .3.

So what should Matt believe about the boundaries of the permissible range? That, it seems, depends on what they actually are. If the lower bound is actually .3, then he should believe that it is between (say) .2 and .4—that is, he should believe that it is *roughly* .3. For even though rational requirements are not wholly transparent to us, they shouldn't be completely opaque to those who reflect carefully. After all, we regard careful reflection on our evidence as valuable precisely because it helps us form rational beliefs—beliefs that better reflect the force of our evidence. If the requirements of rationality were wholly inaccessible, what would justify such a practice?

So if Matt is rational, he believes that the permissible range of credences in *Guilty* spans from *roughly* .3 to *roughly* .7.[6] But if he believes *that*, it would be

[5]See e.g., Christensen (2010), Elga (2013), Horowitz (2014), and Williamson (2000) for sympathetic discussion of similar claims.

[6]Note that this point is not specific to Permissivism. If Epistemic Uniqueness is true, and the unique rational credence in *Guilty* is (say) .5, then Matt is not in a position to know that it is *exactly* .5, and so he too should believe that the rational credence is *roughly* .5. In my reply to the first

irrational for him to adopt credence .3. Why? Because Matt is not certain that .3 is rational, but there are other credences whose rationality Matt does not doubt—he is certain that (say) .5 is rational. But when you are certain that a given credence is rational, it is irrational to adopt any other credence that you are not certain is rational: It is irrational to adopt credences that are, what I'll call, *weakly rationality dominated.*

A bit more precisely. Let an *evidential situation* be a complete specification of which credences are rational responses to one's evidence. Then where c and c' are credences that a subject S might adopt,

> **Weak Rationality Dominance**: c weakly rationality dominates c' for S iff for every evidential situation that S treats as a live option and in which c' is rational, c is too, and in some evidential situation that S treats as a live option, c is rational, but c' is not.

Matt believes that the lower bound is *roughly* .3. To believe that the lower bound is roughly .3 is just to believe that it might be slightly higher or slightly lower than .3. Let's make a simplifying assumption that Matt treats three evidential situations as live options: that the lower bound is .2, that it is .3, and that it is .4. In each of these evidential situations in which it is rational to assign .3 to *Guilty*—i.e., when the lower bound is .2 or .3—it is rational to assign .4 to *Guilty*. But in the evidential situation in which the lowest rational credence is .4, assigning .4 to *Guilty* is rational, and assigning .3 to *Guilty* is not. Hence, in some evidential situation that Matt treats as as a live option, assigning .4 to *Guilty* is rational, but assigning .3 to *Guilty* is not. So assigning .3 to *Guilty* is weakly rationality dominated by assigning .4 to *Guilty* for Matt.

But dominated options aren't rational: It's not rational to adopt a credence that's risky by your own lights when you know of a safer option. It follows that it is not rational for Matt to assign .3 to *Guilty*. But the Permissivist's hypothesis was that it's rationally permissible to adopt *any* credence between .3 and .7. We've now contradicted that assumption. Given what Matt ought to believe about the permissible range, it is irrational for Matt to adopt a credence on the edge of that

objection, I'll show why this does not pose problems for the proponent of Uniqueness. Thanks to an anonymous referee for bringing this to my attention.

range.[7]

The argument generalizes. For any range of putatively permissible credences in a proposition P, it is irrational to assign to P the lowest (highest) value in that range—that is, it is irrational to adopt any credence on the edge of the permissible range.[8] But Permissivism says that there is some range such that any credence in that range is permissible. Contradiction.[9]

(Perhaps you wonder why we should care about rationality dominance: Our ultimate epistemic goal is not rationality *per se* but *accuracy*. Even if we grant this, we ought to recognize that epistemic rationality is a good *guide* to accuracy—in general, more rational credences are more accurate. On this view, we should avoid credences that are rationality dominated because we should do what we can to be most accurate. I return to this issue in my reply to Objection 5.[10])

In the remainder of the paper, I explore various objections to my argument and find them wanting. The upshot is that Permissivism faces a special challenge about the interaction between our first- and higher-order attitudes. Answering the challenge requires taking on new, unattractive commitments about how we form higher-order beliefs or about what it is permissible to do in the face of higher-order

[7] I should note the dominance argument does *not* apply to an extreme version of Subjective Bayesianism that says that we are rationally required to follow the Bayesian formal constraints—probabilistic coherence and conditionalization—but there are no other constraints on what our priors should look like. Why? If any probabilistically coherent prior is permissible, the range of permissible credences in (almost) any proposition will be [0,1]—any credence will be rationally permissible. And if that's right, we can be sure that our credences are rational, and so they won't be weakly rationality dominated. But for many, this extreme version of Subjective Bayesianism will seem *too* permissive. (It doesn't rule out counter-inductivists, for example.) My argument targets any *moderate* Permissivist—that is, any Permissivist who wants to carve out a space between (extreme) Subjective Bayesianism and Uniqueness.

[8] I'm simplifying here. If it is impermissible to take dominated options, then it will be irrational to adopt any credence that is *close enough* to the edge that you believe that it might fall just outside of the permissible range. This means that the argument goes through even if the range of permissible credences has no lowest or highest value.

[9] The Permissivist might object that I oversell the force of the dominance argument. If the range of permissible credences in *Guilty* is narrow, so narrow that Matt is not certain that any particular credence is rational, then his .3 credence in *Guilty* will not be rationality dominated. Granted. But the Permissivist shouldn't rest content. For the guiding motivations for Epistemic Permissivism—e.g., that different people can rationally come to *opposite* conclusions about an issue—strongly suggest that the permissible range of credences will often be very wide. See Rosen (2001) and Schoenfield (2014).

[10] Thanks to an anonymous referee for bringing this to my attention.

uncertainty. Perhaps we will conclude in the end that those commitments are worth carrying to save Permissivism. But if we do accept Permissivism, we should do so with clear eyes. We should know what burdens we shoulder.

1.2 Objections and Replies

Objection 1: *Your argument exploits the vagueness of epistemic permissibility.* It relies on the premise that if a certain credence is not rational, then no credence sufficiently close to it is rational, either. But any premise of that form is soritical, and so ought to be rejected.

Reply: *That* premise is indeed soritical, but my argument doesn't rely on it. I do not say that since .3 is irrational, so too is .31. If my argument did rely on this premise, it would have the (absurd) consequence that there are no rational credences. But it has no such consequence.

To see this, take the limiting case, where there is just one permissible credence in *Guilty*—say, .5. If .5 is the unique rational credence in *Guilty*, then Matt ought to believe that it is *roughly* .5 (reflection doesn't get us all the way to the truth, but it gets us somewhere). Now, if Matt believes that the rational credence is *roughly* .5, then if he assigns .5 to *Guilty*, he won't be certain that his own credence is rational. But that's not *itself* irrational: Although Matt is not certain that his credence is rational, he is not certain, of any particular credence, that it is rational. There is no *other* credence Matt thinks would be better than his own.[11] When the rationality of everything is in doubt, assigning .5 credence to *Guilty* will not be rationality dominated for Matt. Like clear-eyed Permissivists, clear-eyed Uniquers are uncertain about the rationality of their credences; unlike clear-eyed Permissivists, they

[11] Christensen (2010) gives an example of what seems to be a fully rational agent who is uncertain about what the rational credence is. Ava is considering the possibility that the next US President will be a Democrat (D). On page 121, Christensen says, 'Ava gives D some particular crednece, say, .7; this reflects a great deal of her general knowledge, her feel for public opinion, her knowledge of possible candidates, etc. But given the possibility that her credence is affected by wishful thinking, protective pessimism, or just failure to focus on and perfectly integrate an unruly mass of evidence, Ava very much doubts that her credence is exactly what her evidence supports. This seems only natural; indeed, eminently reasonable.'

are not sure that any *other* credences are rational. (By 'clear-eyed Permissivist', I mean someone who believes, of a certain case, that it is a permissive one. By 'clear-eyed Uniquer', I mean someone who believes that there are no permissive cases.)

The dominance argument works against Permissivism because if there is a wide range of credences, Matt can be sure that certain credences in the middle of the range are rational—being in the middle will be safer, by Matt's lights, than being on the edge. But if you think there's just one rational credence, nothing will be perfectly safe by your lights. Matt may assign .5 credence to *Guilty* because this credence will not be rationality dominated by any other credence.

Objection 2: *Ideally rational agents know exactly what rationality permits.* So, in particular, if Matt is ideally rational, he is certain that the permissible range of credences in *Guilty* spans from exactly .3 to exactly .7, and so he is no victim of the dominance argument.

Reply: Maybe.[12] But Epistemic Permissivism is not just a view about *ideally* rational agents, but about ordinary agents like you and me, with all our human limitations. Indeed, Permissivists often tout their view as the only alternative to an objectionably demanding epistemology. Try to imagine yourself in Matt's shoes, they say. You're faced with a mess of evidence. Jones' glove was found on the scene, but another suspect's fingerprints were there too. Three witnesses claim that Jones owned a gun and was prone to violence. Two others deny this. And so on. Could it really be that the only rational response to his evidence is for Matt to become (say) .6453 confident in *Guilty*? Surely not, the Permissivist says. Rationality does not require us to do the impossible—typically there is a wide

[12]Why 'maybe'? Because it isn't clear that even *ideally* rational agents are certain about which credences are licensed by their evidence. Here's one reason to think they aren't. Ideally rational agents aren't always in a position to *know* exactly what they know. In particular, knowledge does not obey negative introspection: Ideally rational agents can fail to know P without also knowing that they don't know P. But if we accept, following Williamson (2000), that our evidence just is our knowledge, then even ideally rational agents won't know exactly what their evidence is—sometimes they will be rationally uncertain about which credences are rational.

range of responses to our evidence that would be rational.[13]

If the Permissivist embraces this motivation, then surely when she says that it is permissible to hold any credence between .3 and .7, she means that it is permissible for someone like Matt—a *non-ideal*, cognitively limited agent—to hold any credence between .3 and .7. But I've argued that this can't be. A non-ideal agent ought not to be certain of the exact boundaries of the permissible range, so he cannot adopt a credence on the edge of that range.

But even if the Permissivist eschews this motivation, I don't think she's much better off. To be sure, the dominance argument does not apply to those who are certain of the exact boundaries of the permissible range. So, for any credence between .3 and .7 in *Guilty*, perhaps there is an ideally rational agent who holds that credence. But the argument still applies to ordinary agents—*their* options will be much more constrained. This leaves us with quite a surprising account of rationality, one that affords ideally rational agents many, many options, and limits ordinary agents to just one.

This is exactly the *opposite* of what we should expect from a Permissivist who recognizes a distinction between ideal and non-ideal rationality. Permissivism is populist epistemology, a view for ordinary folk. Impermissivists ignore the realities of our actual cognitive lives. Permissivists don't. If the Permissivist grants that any form of rationality is impermissive, it should be an idealized notion, one that abstracts away from our actual cognitive limitations, and so, by their lights, has little relevance for people like us. Absent a suitable story about why things would be reversed—why it would be ideal rationality that is permissive and non-ideal rationality that is impermissive—we should be suspicious of an appeal to ideal rationality as a way of evading the dominance argument.

Objection 3: *Matt should believe that he is in the middle of the permissible range.* Suppose we grant that it is impermissible to adopt dominated options. Then we must say that, *if he believes that the boundary of the permissible range is roughly .3*, it is irrational for Matt to be .3 confident in *Guilty*. Matt shouldn't regard himself as taking an unnecessary risk. But that doesn't mean the .3 credence has to go, as I've argued. Rather, it's Matt's higher-order beliefs that should change: Matt

[13] See e.g., Schoenfield (2014) for arguments like this.

should believe that his .3 credence is close to the *middle* of the range of permissible credences in *Guilty*, so that he is certain that .3 is rational.

Reply: If you take this strategy, you need a story about how we form higher-order beliefs that explains why agents on the edge of the permissible range must always believe falsely that they are (roughly) in the middle. To be sure, agents can form false beliefs about what rationality permits if they have *misleading* higher-order evidence. Matt may get misleading higher-order evidence suggesting that the permissible range of credences in *Guilty* spans from roughly .1 to roughly .5. I don't doubt that it would then be permissible for him to assign .3 to *Guilty*. But Epistemic Permissivism is not just a view about what we are permitted to believe when we are *misled*. We need an account of how we form higher-order beliefs that explains how it can be rational for agents on the edge to believe falsely that they are in the middle, even in the absence of misleading higher-order evidence.

The kind of story about higher-order belief formation that we are interested in is one that tells Matt and Abby to believe that they are in the middle of the permissible range. So, since Matt assigns .3 to *Guilty*, he ought to believe that the range of permissible credences in *Guilty* is roughly .1-.5, and since Abby assigns .7 to *Guilty*, she ought to believe that the range of permissible credences is roughly .5-.9. What kind of method must Matt and Abby be using to form these higher-order beliefs? It must be one that takes into account their *own* credences in *Guilty*.

Perhaps the story goes like this. In general, you are rationally entitled to believe that your own credences and beliefs are rational. So, when Matt examines the evidence and adopts .3, he is thereby permitted to infer that .3 is rational. Similarly, when Abby examines the evidence and adopts .7 credence, she is thereby permitted to believe that .7 is rational. Now, the story can't stop there, for since Matt and Abby are clear-eyed Permissivists, they think other credences are rational, too. Which ones? A natural thought: Those that are sufficiently close to their own! Matt reasons, 'I know that .3 is rational. So any credence as low as .1 or as high as .5 is probably rational, too', and Abby reasons, 'I know that .7 is rational, so anything as low as .5 and as high as .9 is rational, too.' Can this procedure explain how Matt and Abby could *rationally* believe that they are in the middle of the range?

I think not. For if your credence is sufficiently close to the edge, this procedure will lead you astray about the upper and lower bounds of the permissible range, as it does for Matt and Abby. Now, the unreliability of the method is not itself the problem—perhaps we are sometimes rational to use unreliable methods. What's worrying is that a clear-eyed Permissivist is in a position to *know* that the method is unreliable.

To see why, it will be instructive to compare Permissivism to Uniqueness on this issue. Suppose that I am a clear-eyed Uniquer. In that case, I know that, if my first-order belief or credence is rational, then if I come to believe that my own credence is rational, my belief will be true. In short, if Uniqueness is true, then the rational entitlement principle guarantees that *rational* agents will form true beliefs about rationality, and clear-eyed Uniquers are in a position to rationally believe this about themselves.

Things are different for Permissivism. Matt is a clear-eyed Permissivist. He believes that his credence in *Guilty* is rational and infers that all credences close to his are rational. That's how Matt comes to believe that the range of permissible credences in *Guilty* is roughly .1 to roughly .5. But although being rational guarantees that you will form true beliefs about the rationality of your *own* credence, it does not guarantee that you will form true beliefs about the upper and lower bounds of the permissible range of credences—it doesn't guarantee that you will form true beliefs about what other credences are rational.

But Matt in a position to *know* this about the method he is using. Matt should be highly confident that he is on the edge of the permissible range of credences in *some* propositions, and for all he knows, *Guilty* is one of them. So, Matt should recognize that, for all he knows, his credence is on the edge of the range of permissible credences in *Guilty*. Matt should doubt the conclusion of the method he is using: He should think, 'Even if my credence in *Guilty* is rational, I might be on the edge of the permissible range, in which case my belief that all credences close to my own are rational will be false!' But it can't be rational to use a method whose conclusions one doubts. If I am in a position to know *a priori* that a conclusion I've drawn might be false, I shouldn't believe it. Pending a suitable story about higher-order belief formation, we ought to be suspicious of a view that says that we must always believe that our credences are close to the middle of the permissible range.

Objection 4: *You assume that if Permissivism is true, then Matt can rationally believe he is in a permissive case.* But some Permissivists deny this. For example, Stewart Cohen (2013) is a Permissivist who defends **Doxastic Uniqueness**, the claim that a subject cannot rationally hold one credence while believing that some other credence is just as rational. That is, a subject can never rationally believe that she is in a permissive case. If Matt doesn't believe that there is a range of permissible credences in *Guilty*, the threat of dominance evaporates.

Reply: Granted. But the traditional motivations for Permissivism strongly suggest that if Permissivism is true, then we can, at least sometimes, rationally believe that we are in a permissive case.[14]

Permissivist usually motivate their view, at least in part, by reflection on particular cases—cases of disagreement among jurors about whether Jones is guilty, among paleontologists about what killed off the dinosaurs, and among philosophers about whether we have free will. But if we can know that a particular case is permissive when we're doing epistemology, what could stop us from continuing to know that a particular case is permissive *when we're in one*?

A second way of motivating Permissivism, the 'competing theoretical virtues' argument, also suggests that we often know that we're in a permissive case. Permissivists say that what it's rational for us to believe depends not just on the content of our evidence, but on how we balance certain theoretical virtues against each other—things like simplicity, predictive strength, and explanatory power. There are many different, yet equally rational ways of balancing these virtues against one another, and they will often yield different levels of confidence in various hypotheses.[15]

But surely we can tell, at least sometimes, how simple, predictive, and explanatory a certain hypothesis is. If we can, then we can also determine, say, that balancing simplicity and predictive power in *this* way would yield high credence in *Guilty*, and weighing them in *that* way would yield lower credence in *Guilty*.

[14]See Ballantyne and Coffman (2011), Douven (2009), Kelly (2013), Schoenfield (2014), and Titelbaum and Kopec (2015) for defenses of clear-eyed Permissivism, the view that Permissivism true and we're sometimes in a position to know, of a certain case, that it is permissive.

[15]See, for example, Douven (2009), Schoenfield (2014), and Titelbaum (2015a) for arguments along these lines.

But if we also know that both of these ways of balancing the theoretical virtues are rational, we can put two and two together, and conclude that high credence and low credence in *Guilty* are both rational.[16]

Objection 5: *It can be permissible for Matt to hold onto a credence that is rationality dominated.* Matt's credence has something else going for it: He expects it to be most accurate. Why? Because Matt's credence is recommended by his own epistemic standards, which he *endorses*: Matt expects the credences recommended by his standards to be more accurate than those recommended by any other standards.

This argument is not new: It's Schoenfield's (2014) response to White's (2005) charge that if Permissivism is true, then it should be fine to arbitrarily switch from one permissible credence to another.[17] We can think of Schoenfield's brand of Permissivism as a kind of **Intrapersonal Epistemic Uniqueness**: There are many, equally permissible sets of epistemic standards, but once you've settled on one, you have reason to adopt the credences your standards recommend. (For Schoenfield, that's because if you're rational, you expect your standards to maximize expected accuracy.) [18] So, *contra* White, it is never rational to arbitrarily switch from one permissible credence to another.

If we're convinced by Schoenfield's reply to White's, might we use it to reply to my challenge as well?

Reply: There are two ways of understanding Schoenfield's reply on behalf of the Intrapersonal Epistemic Uniquer, one weaker, one stronger: the weaker, though plausibly an effective response to White's challenge, is no objection to mine; the stronger is indeed an objection to the dominance principle, but, intuitively, it is far too strong.[19]

The weaker version of Schoenfield's reply says that expecting some credence to be most accurate is *sometimes* a reason to prefer it—in particular, it is a reason

[16]Thanks to an anonymous referee for suggesting this point.

[17]White presents many arguments against Permissivism, but I take this to be the central objection unifying all of them.

[18]Kelly (2014), Meacham (2014), and Titelbaum and Kopec (2015) also defend Intrapersonal Uniqueness, and discuss how this view helps us respond to White.

[19]Thanks to Bernhard Salow for discussion on the arguments of this section.

to prefer it when the rationality of the various credences you're considering *is not in doubt*. This is the thought that in fact motivates Schoenfield's reply to White above (see footnote 21). Plausibly, Schoenfield has answered White's challenge—she has explained why, when you know that some other credence is just as rational as yours, you still have reason to prefer your own.

But the weaker version of Schoenfield's reply does not help the Permissivist answer *my* challenge—she does not (nor does she intend to) explain why it would be permissible to hold onto credences that are rationality dominated. After all, cases in which your credence is rationality dominated are precisely those cases in which the rationality of one of the credences you are considering *is* in doubt.

The stronger version of Schoenfield's reply says that expecting your credence to be most accurate is *always* a reason to prefer it—in particular, it is a reason to prefer it even when the rationality of that credence *is in doubt*. This is is indeed an objection to the dominance principle, the principle that it is always irrational to adopt weakly dominated credences. The stronger version of Schoenfield's reply says that you should stick to your credence when it is rationality dominated because you expect it to be most accurate.

But it's not an objection the Permissivist should be happy to pursue. It implies that we should never be moved to revise our credences by evidence that we've been less than fully rational; instead we should level-split—e.g., we should both remain highly confident that Jones is not guilty *and* believe that our evidence supports lower confidence in this proposition. This is not a welcome consequence. Those who refuse to revise their beliefs in the face of evidence that they are irrational seem over-confident, indeed dogmatic—they seem paradigmatically irrational.

Suppose Jill has examined all of the evidence and becomes highly confident that the Warriors, her favorite NBA team, will win the championship. A trusted friend tells her that she always overestimates the likelihood of favorable outcomes. The Permissivist we're considering says that Jill should remain highly confident that the Warriors will win despite her friend's warning. Since she expects her high credence to be most accurate, she needn't be worried by evidence that she's been irrational.

This doesn't seem right. When she has reason to believe her credence is irrational, Jill shouldn't be able to appeal to the perceived accuracy of her credence as a reason to hold onto it. Compare this to a case of all-out beliefs. Suppose that Jill simply believes that the Warriors will win, and her friend tells that the evidence doesn't support such high confidence. Jill couldn't reply to her friend's concern: 'Well, I must have gotten lucky—even though the evidence supports lower confidence, I've wound up with a true belief!'[20] Return to the case of credences. Jill can't respond to her friend's concern about the rationality of her credence with: 'Well, I must have gotten lucky—even though my evidence supports a lower credence, my credence is more accurate!' Evidence that she has been irrational should make her doubt the accuracy of her credences and beliefs, and she should lower her confidence accordingly.

Similar things can be said of Matt's credence in *Guilty*. He worries that his credence might be too low. But he's sure that it's not too high—he's certain that it would be rational to assign .4 to *Guilty*. The Permissivist we're considering says that he needn't be moved by doubts about the rationality of his .3 credence in *Guilty* because he expects it to be most accurate. But again, this seems wrong. Matt cannot appeal to the perceived accuracy of his credence as a reason to maintain that credence when its rationality is in doubt.[21]

This is connected to a worry I mentioned earlier: Don't rational agents care rationality *only as a guide to accuracy*, and not for its own sake? If so, accuracy always comes first. But then my dominance principle is false: Because he expects his credence to be more accurate than any other credence, Matt should stick with .3 even if he isn't sure it is rational and he is sure that (say) .4 is rational. To shift from .3 to .4 in order to ensure rationality—at the cost of accuracy by Matt's

[20]For defenses or sympathetic discussion of the claim that it is irrational to maintain some credence or belief when you acquire evidence that it is irrational, see Christensen (2010), Elga (2013), Greco (2014), Horowitz (2014), Sliwa and Horowitz (2015), Smithies (2012), and Titelbaum (2015b). For criticism see Coates (2012), Lasonen-Aarnio (2014), Lasonen-Aarnio (2015), and Williamson (2011).

[21]It is Schoenfield's discussion of irrelevant influences on belief that makes explicit that she endorses only the weaker version of Intrapersonal Uniqueness. She's interested in cases where you learn that your belief was influenced by an irrelevant factor. In cases like this, it seems we should revise our earlier beliefs. Schoenfield explains this intuition by saying that we should revise our beliefs *when we have reason to believe that they are irrational.* (See pages 203-206.) Schoenfield accepts that if you have evidence that you've picked standards in a way that was unlikely to leave you with rational ones, then you ought to change them. It's a short step from this thought to my dominance principle.

lights—would be irrational.

But I deny that it can be rational for Matt to expect his credence to be most accurate when he has reason to believe that his credence might be irrational, and he is sure that some other credence is rational. If Matt were sure that his credence is rational, then perhaps it would be rational for him to expect it to be most accurate. But the moment he starts doubting the rationality of his credence, continuing to assume that it is most accurate seems overly self-confident and dogmatic.

Even if we don't care about rationality for its own sake, it *is* a good guide to accuracy—in general, more rational credences are more accurate than less rational ones. But then we ought to see it that way—if we expect one credence to be more rational than another, we should also expect it to be more accurate, and so we should prefer it. Since Matt expects .4 to be more rational than .3—he's sure that .4 is rational but he is not sure that .3 is—he should also expect .4 to be more accurate than .3, and Matt should revise his credence accordingly. On this view, avoiding dominated credences is just part of doing what we can to be most accurate.

Let's take a step back. As we've seen, I'm not the first to object to Permissivism. White argued that if you know that you're in a permissive case, then it is okay to arbitrarily switch credences. As we saw in Objection 4, some Permissivists simply deny that we ever know that we are in a permissive case, and such Permissivists escape my dominance argument.

But the Permissivists who endorse the traditional motivations for the view—e.g., competing theoretical virtues—say that you *can* know that you are in a permissive case, but it is nonetheless not okay to arbitrarily switch. These Permissivists must say that rational agents have some to reason privilege their own credences even when they know that some other credence is just as rational. Perhaps that reason is that the standards of a rational agent maximize expected accuracy for that agent, as Schoenfield argues. Or perhaps it's simply that we should be (diachronically) consistent; we shouldn't change our minds arbitrarily.[22]

But whatever the reason is, we must ask the question that I asked in response

[22]See Titelbaum and Kopec (2015) for a discussion of how appealing to certain norms of diachronic consistency helps us reply to White.

to Schoenfield: What is the *force* of the reason? Do we *always* have reason to adopt the credences that our standards recommend? Or rather, can it be *defeated* by evidence that our credences (and so our standards, too) are irrational? As we've seen, the Permissivist should accept that if you have reason to believe that your standards are irrational, then whatever reason you had to adhere to your standards is defeated—when the rationality of a certain credence is in doubt, that it is recommended by *your* standards is no reason to prefer it.

It is a far cry from the thought that you shouldn't change your mind arbitrarily to the thought that you shouldn't change your mind *even when you have reason to believe that your present attitude is irrational.* But I have argued that *this* is the predicament of the clear-eyed Permissivist on the edge of the permissible range. It's not that he believes that some other credence is just as rational as his. No, it's that he expects some other credence to be *more* rational than his. (Matt is sure that .5 is rational but he is not sure that .3 is rational.) To privilege your credence when the rationality of that credence is not in doubt is one thing; to do so when you have reason to believe it might be irrational, and you're sure that some other credence is rational, is quite another.

The dominance argument goes through so long as you admit that it is rational to revise your credences when you're not sure that your credence is rational but you *are* sure that some other credence is rational. And *any* Permissivist should admit this much.

1.3 Conclusion

I've argued that if you are on the edge of the permissible range of credences in a certain proposition, your credence will be rationality dominated by certain credences closer to the middle. Since dominated options aren't rational, it's not rational to adopt a credence on the edge of the permissible range. But Permissivism says that, for some such range, *any* credence in that range is rationally permissible. I have considered some objections to my argument and found them wanting. Permissivism, in its traditional form, cannot be right.

Chapter 2

Belief and Probability

We believe lots of things. Many of those beliefs are rational. I rationally believe that there's a bright screen in front of me, that I didn't have breakfast this morning, that two plus two equals four. If I'm rational to believe these things, what grounds this fact? What makes it rational for us to believe certain propositions? And what is it for the rationality of our beliefs to be defeated? Many epistemologists will want answers to these questions. They will want to reduce rational belief to something else. But there's a certain widely influential character in contemporary epistemology—the *Epistemic Bayesian*, as I'll call him—who will insist that only one sort of reduction will do: a reduction to subjective probability.[1]

The Epistemic Bayesian thinks that the opinions of a rational subject, at a given time, can be faithfully modeled by a probability function, which assigns, to each proposition, a precise probability.[2] To his mind, all varieties of rational opinion can be—must be, somehow—reduced to facts about these subjective probabilities. If I am entitled to believe that I skipped breakfast, there is a probabilistic explanation for that fact; if my belief is defeated, there is a probabilistic explanation for that too.

Here I show that rational belief resists probabilistic reduction. There's a way

[1]I borrow the term 'Epistemic Bayesian' from Joyce (2005).

[2]More exactly, this is what the *Precise* Epistemic Bayesian thinks. As we'll see, there are also *Imprecise* Epistemic Bayesians. They propose to model rational subjects with a *set* of probability functions, a kind of credal committee. I return to this distinction in §2. The term 'Epistemic Bayesian' applies equally to the precise and imprecise varieties.

to rationally lose belief in a certain hypothesis without losing confidence—a kind of *non-probabilistic defeat.* The problem stems from conflicting evidence, the sort of evidence you have when two sources of information contradict each other, such as when doctors offer conflicting diagnoses, or when film critics you trust write discordant reviews. Under the right circumstances, conflicting evidence can defeat your entitlement to believe a certain hypothesis without disconfirming it—that is, without making it any less probable. I present a case of conflicting evidence that illustrates this phenomenon. I consider three probabilistic theories of rational belief—one simple, two more sophisticated. None of these theories can explain why it's rational to stop believing in the cases I describe.

I can't prove that no adequate probabilistic theory of rational belief will ever be devised. But I can give reasons for pessimism—reasons to doubt that probabilistic theories have the *kinds* of resources we need.

2.1 Who is the Epistemic Bayesian?

Epistemic Bayesianism is the conjunction of two theses: *Bayesianism* and *Lockean supervenience*.

Bayesianism is built on the notion of a *credence* or degree of confidence. We don't just believe and disbelieve; we believe and disbelieve to varying *degrees*. To model our degrees of belief, Bayesians say that each of us is equipped with a credence function Pr (in the case of *Precise Bayesianism*) or set of credence functions (in the case of *Imprecise Bayesianism*) from propositions to real numbers between zero and one. (To simplify things, I'll introduce Precise Bayesianism now, and Imprecise Bayesianism later, but it's important to remember that the Epistemic Bayesian may endorse either doctrine.) $Pr(P)$ represents your degree of confidence (your credence) that P is true.

The Bayesian says that if you're rational, then your credence function will also be a *probability function*. This means that it obeys the axioms of the probability calculus, listed below:[3]

[3]These are often referred to as *Kolmogorov's axioms*, after the Russian mathematician Andrey Kolmogorov, who was the first to articulate these axioms as the foundation of probability theory. See Kolmogorov (1933).

(1) *Non-Negativity*. For any proposition P, $Pr(P) \geq 0$

(2) *Normality*. For any tautology T, $Pr(\mathrm{T}) = 1$

(3) *Finite Additivity*. For any mutually exclusive propositions P and Q, $Pr(P \vee Q) = Pr(P) + Pr(Q)$.

The probability axioms are the first three requirements of Bayesianism—the constraints that a subject's (unconditional) credence function must satisfy at any given time in order to be rational. (When a subject's credence function Pr is a probability function, I will often refer to her credences as *subjective probabilities*.)

Bayesians insist that subjects also have various *conditional* credences. $Pr(P|Q)$ represents one's credence in P conditional on Q—one's credence that P is true on the supposition that Q is. Conditional credences are related to unconditional credences according to the *Ratio Formula*:

(4) *Ratio Formula*. Where $Pr(Q) > 0$, $Pr(P|Q) = Pr(P \wedge Q)/Pr(Q)$.

(If $Pr(Q) = 0$, $Pr(P|Q)$ is undefined.)

Finally, Bayesians use conditional credences to state a rule governing how rational subjects update their beliefs:

(5) *Conditionalization*. For any time t_1, and any later time t_2, if you learn all and only Q between t_1 and t_2, then for any proposition P, $Pr_{t_2}(P) = Pr_{t_1}(P \mid Q)$.

Informally speaking, Conditionalization says that your credence in P at t_2 upon *learning* Q should equal your prior t_1 credence in P had you merely been *supposing* Q. The Bayesian says that if you're rational, you update your credences by conditionalizing on your evidence. Conditionalization is the final requirement of Bayesianism.[4]

[4]There are a variety of arguments that have been given in support of Conditionalization. See, for example, Teller (1976) for a *Dutch Book* argument for Conditionalization—an argument that if you don't conditionalize, a Dutch Book can be made against you. (Teller credits David Lewis for the argument.) See Williams (1980) for an argument that appeals to the Principle of Minimum Information, van Fraassen (1999) for an argument that appeals to the Principle of Reflection, and Greaves and Wallace (2006) for an argument that Conditionalization maximizes expected accuracy.

We've introduced Bayesianism, the first thesis of Epistemic Bayesianism. The second thesis is what I call *Lockean Supervenience*:

> *Lockean Supervenience.* For any time t, your rational unconditional and conditional credences at t determine (i) what you should believe at t, and (ii) what you should believe as you gather more information.

Lockean Supervenience says that facts about rational belief supervene on facts about rational credences. There can be no difference in what it is rational for you to believe without some difference in which credences it is rational for you to have. Or, put the other way around, once we've settled which credences it is rational for you to have, we've settled what it's rational for you to believe.[5]

As I understand him, the Epistemic Bayesian is not an *eliminativist*. He's happy to talk about belief and rational belief. But he is an adamant *reductionist*. The *fundamental* epistemic facts are facts about rational probabilities, and all other facts about rational opinion are to be explained in terms of these fundamental probabilistic facts. If it's rational for me to believe there is a bright screen before me, there is a purely probabilistic explanation for that fact; if the rationality of my belief is defeated, there is a purely probabilistic explanation for that too.

In what follows, I argue against Epistemic Bayesianism. More specifically, I argue that if rational credences are what the Bayesian says they are, then Lockean Supervenience is in trouble. In certain circumstances, you can learn something that defeats your entitlement to believe a certain hypothesis without making it any less probable. If beliefs can change when subjective probabilities don't, rational belief isn't tethered to the logic of subjective probability. Belief is irreducible to subjective probability.[6]

[5]Some version of Lockean Supervenience is assumed either implicitly or explicitly by many epistemologists and philosophers of science working in the Bayesian tradition. See, among others, Clarke (2013), Dorst (forthcoming), Foley (2009), Leitgeb (2013, 2014, 2017), Sturgeon (2008), and Weatherson (2005) for expressions of this Lockean sentiment.

[6]I am not the first to object to Lockean Supervenience. Buchak (2014) criticizes the Lockean Thesis (roughly, the claim that you believe P just in case you are sufficiently confident in P) on the basis of what she calls *merely statistical evidence.* Suppose, for example, that a car is hit in the night by a bus. There are two bus companies in town: Blue Bus owns 95% of the buses, while Red Bus owns 5%. If you believe that Blue Bus did it, you should fine them. But you can't fine them on the basis of these statistics! So you don't believe that Blue Bus did it, even though it is 95% likely that

2.2 A Case

One of your friends needs a kidney transplant. Adam, Ollie, and Arnie are candidate donors. Each will undergo state-of-the-art blood testing. Before testing, each tells you his blood type.[7]

> **Kidney Transplant (Stage One).** Adam says that he is type A. Arnie says that he is type A. Ollie says that he is type O.

People tend not to misreport their own blood types. You believe what you've been told—that Adam and Arnie are type A, and that Ollie is type O—and you are rational in so doing.

The test says whether two people are the same blood type, but it doesn't tell you what type they are. Ollie and Arnie are tested first.

> **Kidney Transplant (Stage Two).** The test reveals that Arnie and Ollie are the same type.

The test is always right. When it says that Arnie and Ollie are the same blood type, you trust it, and conclude that they're either both type O or both type A. You trust Arnie and Ollie equally, we'll suppose. Since you have no reason to trust Arnie over Ollie, you don't believe they are both type A. And since you have no reason to trust Ollie over Arnie, you don't believe they are both type O. Still, you should continue to believe that Adam is type A. Discovering that one of Ollie and

they did. While I agree with Buchak that belief doesn't reduce to credence, here's a worry about her reasons for rejecting the Lockean Thesis. Hawthorne et al. (2016) persuasively argue that the norms for belief are *weak*—in particular, they are weaker than the norms for certainty or for knowledge. I can say 'I believe it's raining, but I'm not sure' or 'I believe it's raining, but it might not be.' I *cannot* say, 'I'm sure it's raining, but it might not be.' According to how we ordinarily talk, believing P is compatible with believing that P might be false, and we often *do* believe on the basis of merely statistical evidence: I believe that the coin won't land heads 100 times in a row, even though I have nothing but statistical evidence to go on. In response to these sorts of considerations, some will say (Greco (2016) explicitly does) that they aren't talking about the notion picked out by the natural language 'belief', but a technical notion akin to certainty. I have no objection to them, but their topic is not mine.

[7] The structure of this example was inspired by an example originally given by Matthew Ginsberg (1986) that was intended to raise trouble for certain principles in the logic of counterfactuals.

Arnie was mistaken is not cause to question Adam's testimony.[8]

Finally, Adam and Ollie undergo testing.

> **Kidney Transplant (Stage Three).** The test reveals that Adam and Ollie are the same blood type.

You should now be sure that Ollie, Arnie, and Adam are the same blood type. Should you believe that they are all type O? Plainly not. Adam and Arnie both said that they are type A. Should you believe they are all type A? That doesn't seem right, either. Ollie said he is type O. Perhaps you're *more confident* that they're all A than that they're all O—you have two pieces of evidence suggesting A and just one suggesting O. But to *believe* that Adam, Arnie, and Ollie are type A would be to conclude that it is Ollie who made the mistake. And, intuitively, you are not entitled to draw that conclusion. You should leave open that Ollie is right, and that Adam, Arnie, and Ollie are all type O.[9]

Here's what one would *like* to say about Kidney Transplant. At Stage One, you are entitled to rely on Adam's testimony about his blood type, entitled to believe Adam is A. Typically people accurately report their blood types, and absent evidence to the contrary, you should assume Adam is no exception. When you learn that Adam is the same type as Ollie and Arnie at Stage Three, you are no longer entitled to trust Adam. Trusting Adam now means *siding* with Adam (and Arnie), and you aren't entitled to do *that*—trusting Adam *over* Ollie wasn't part of the deal. Your entitlement was provisional and circumscribed. It said that *absent evidence to the contrary*, you're entitled to assume Adam knows his own blood type. But

[8]Normally when you learn someone was wrong about his blood type, you needn't worry that everyone else is. To be sure, there are special circumstances in which learning that one person is wrong is evidence of widespread error—say, if you know that blood tests are either perfectly reliable or *anti*-reliable, biased away from the truth. But we can just stipulate that this case is not like that.

[9]Here's a way to underscore the intuition. Simplify the case so that Arnie, Adam, and Ollie are tested all at once, and the test reveals that all three have the same blood type. Plainly, it would be irrational to outright conclude that they're all type A. Instead, you should leave open that they are all O. But the only difference between the simplified case and the original case is in *how long* it takes for you to learn that they all have the same blood type. In the simplied case, you learn all at once that Arnie, Adam, and Ollie have the same blood type. In the original case, you learn this in stages—first you learn that Arnie and Ollie are the same, then you learn that Adam and Ollie are the same. Surely this difference in timing shouldn't affect what you believe. Your *evidence* about Adam, Arnie, and Ollie's respective blood types is exactly the same in the two cases.

now you you *have* evidence to the contrary. You are no longer entitled to believe Adam is A.

That's what one would like to say, anyway. But we've said it in a language foreign to probabilistic theories of evidence and rational opinion. Can we explain, in probabilistic terms, why you're entitled to believe that Adam is A at Stage One and no longer entitled to believe this at Stage Three? This is what the Epistemic Bayesian demands. We should certainly hope we can. Probabilistic theories of evidence are precise, tractable, and predictive, promising deep and unifying explanations of diverse epistemic phenomena. Yet I'll show that cases like Kidney Transplant pose a serious challenge for probabilistic reductions of rational belief. The simplest probabilistic theory predicts no difference in what you are entitled to believe between Stage One and Stage Three, and more sophisticated probabilistic theories fare no better.

2.3 The Simple Theory

We said that you are entitled to believe that Adam is A at Stage One and Stage Two of Kidney Transplant, and no longer entitled to believe this at Stage Three: learning that Adam is the same blood type as Ollie (and Arnie) defeats your belief. The Epistemic Bayesian must provide a probabilistic explanation for this change. The simplest idea is to posit a threshold for rational belief. It's rational to believe P just in case your rational credence in P exceeds that threshold.

> **The Simple Theory**. There is some threshold t such that it's rational for you to believe P if and only if your rational credence in P exceeds t.

I assume that t is greater than .5 (it's never rational to believe P and $\neg P$) and less than 1 (belief can't require absolute *certainty*, lest we lose too many of our beliefs).

Here's how the story would go. It's rational to believe that Adam is A at Stage One and Stage Two because it is sufficiently probable that Adam *is* A on your evidence at those times. It's not rational to believe that Adam is A at Stage Three because, given your evidence then, it's not sufficiently probable that Adam is A. In short, learning that Adam is the same blood type as Ollie (and Arnie) disconfirms the hypothesis that Adam is blood type A. It's more likely that Adam is A on your evidence at Stage One and Stage Two than it is on your evidence at Stage Three. That's why you lose your entitlement to believe.

2.3.1 Against The Simple Theory

The Simple Theory won't do. We can *prove* that you should be no less confident that Adam is type A at Stage Three than you are at Stage One and Stage Two. To do so, we need three assumptions.

The first assumption is that, contrary to fact, the population is divided equally between type A and type O, and you're sure of this. You begin with no other relevant information about Ollie's blood type, about Adam's blood type, or about Arnie's blood type. So before Ollie tells you his blood type, you are 50% confident that Ollie is O and 50% confident that Ollie is A. The same goes for Arnie and Adam.

Second, I assume you trust Ollie and Arnie equally. This means that your credence that Arnie is A conditional on Arnie saying that he's A is equal to your credence that Ollie is O conditional on Ollie saying that he's O. Given our first assumption that you start out 50%-50% about Ollie's blood type, and 50%-50% about Arnie's blood type, it follows that when you learn that Arnie and Ollie are the same type, you are no more confident that one is right than that the other is. (See §4.2 for a longer discussion of why this is so.) At Stage Two, you are 50% confident that Arnie and Ollie are both type A and 50% confident that they are both type O.

Finally, I assume that whether Adam is type A is *probabilistically independent* of whether Ollie is type O at Stage Two: learning that Ollie is type O would not increase your confidence that Adam is type O (and *vice versa*); similarly, learning that Ollie is type A wouldn't increase your confidence that Adam is A (and *vice versa*).[10] In short, learning one of their blood types shouldn't change your attitude about the other's.

We said that you're entitled to believe that Adam is type A at Stage Two on the

[10]Importantly, probabilistic independence is only required at Stage Two. The proposition that Adam is A is clearly not probabilistically independent of the proposition that Ollie is A at Stage Three. By then you are sure that Adam is A just in case Ollie is, so learning either of their blood types would be sure to change your opinion about the other's blood type.

basis of Adam's testimony. You have conflicting information about Ollie—Ollie says that he (Ollie) is type O, but Arnie says he (Arnie) is type A, and you know they're the same. You are 50%-50% about Ollie's blood type at Stage Two. At Stage Three you learn that Adam and Ollie are the same blood type, and you're *no longer entitled to believe* that Adam is type A. The Simple Theory says that you lose your entitlement to believe that Adam is A because the proposition that Adam is the same as Ollie disconfirms the hypothesis that Adam is A. The probability that Adam is A conditional on Adam and Ollie being the same is lower than the unconditional probability that Adam is A at Stage Two.

The trouble is that it follows from our three assumptions and the probability calculus that the probability that Adam is A conditional on *Adam and Ollie are the same type* is the same as the unconditional probability that Adam is A at Stage Two. In the next section, I explain why.

2.3.2 Explanation and Diagnosis

Remember what we would like to say about Kidney Transplant. At Stage One, you are entitled to believe that Adam is A on the basis of his testimony. Learning that Adam is the same type as Ollie (and Arnie) defeats your entitlement. You are not entitled to side with Adam (and Arnie) over Ollie. I said that this is not something we can say if the Simple Theory is true. But it would be well to explain why.

Here's an attempt. You think that Ollie and Arnie are equally reliable. You are just as confident that Ollie is right about his blood type as you are that Arnie is right about *his* blood type. Ollie's testimony confirms the hypothesis that Ollie and Arnie are type O to the same degree as Arnie's testimony disconfirms it. Their conflicting reports cancel each other out: taken together, they have no effect on the probability that Ollie is O or on the probability that Arnie is A. When you learn of the conflict, you should return to the credences you had before Ollie or Arnie said anything—50%-50% about Ollie, and 50%-50% about Arnie. (Of course there is one difference—now you know Ollie and Arnie are the same type.) Probabilistically speaking, it's as though you never spoke to Ollie or Arnie; all you know is that half the population is O and half is A. When you discover that Adam is the same type as Ollie, we should think of your predicament as one in which you have just this statistical information and Adam's testimony to go on. But we've already said what you're entitled believe when *that's* your evidence. You're entitled to trust Adam's testimony and believe that Adam is A.

We can turn these informal remarks into a more rigorous explanation composed of two claims. The first is that Ollie's and Arnie's conflicting reports cancel each other out, leaving you 50%-50% about Ollie. Here's one way to see why that should be true. Ollie and Arnie are just as likely to disagree about their shared blood type when Ollie is right (and they're both type O) as when Arnie is right (and they're both type A). Formally:

(1) $Pr(\textit{conflicting reports}|Ol\ O) = Pr(\textit{conflicting reports}|Ol\ A)$

From (1) we can infer that the conflicting reports don't affect the net balance of your evidence with respect to the hypothesis that Ollie is type O. Taken together, they make it no more likely that Ollie is type O than that he is type A, and *vice versa*. In other words:

(2) $Pr(Ol\ O|\textit{conflicting reports}) = Pr(Ol\ A|\textit{conflicting reports})$.

And finally, from (2) it follows that the probability that Ollie is O, on the supposition that Ollie and Arnie issue conflicting reports about their shared blood type, is equal to the probability that Ollie is O *prior* to any reports being issued. Learning of the conflicting reports should not change your credence that Ollie is O. Formally:

(3) $Pr(Ol\ O|\textit{conflicting reports}) = Pr(Ol\ O)$.

Since you are 50%-50% about Ollie's blood type before you hear from Ollie or Arnie, you are still 50%-50% after.

The second claim is that, because you are 50%-50% about Ollie's blood type, when you learn that *Adam* and Ollie are the same blood type, it's as though you have just Adam's testimony to go on. Your evidence about Ollie is probabilistically idle; it ought to have no effect on your credence that Adam is A. This claim is underwritten by the following fact:

No Change. For any propositions P and Q, if P and Q are probabilistically independent relative to Pr, and $Pr(Q) = .5$, $Pr(P) = Pr(P|P \equiv Q)$.[11]

No Change says that if P is probabilistically independent of Q, and you are 50% confident in Q, then learning $P \equiv Q$ should not change your opinion about P. If you are confident in P, you must remain confident upon learning $P \equiv Q$ (and raise your confidence in Q accordingly).

No Change is the central principle at work in our proof. Consider Stage Two of Kidney Transplant. We just said you are 50% confident that Ollie is A. We've also said that whether Adam is A is probabilistically independent of whether Ollie is. It follows from No Change that the probability that Adam is A, on your current evidence, is equal to the probability that Adam is A, *on the supposition that Adam and Ollie are the same blood type*. That Adam and Ollie are the same type does not disconfirm the hypothesis that Adam is A.

The Simple Theory says that learning that Adam is the same type as Ollie (and Arnie) defeats your entitlement to believe that Adam is type A because it disconfirms that hypothesis. We've now seen that this simply isn't so. The fact that Adam is the same type as Ollie (and Arnie) makes it no more or less likely that Adam is A.[12]

To get a better sense of the problem, consider a variation on Kidney Transplant. Adam reports that he is blood type A, and you believe him. This time there's been no word from Ollie or Arnie. Neither has told you his blood type. You know that half the population is type O, but have no other relevant information about Ollie's blood type. So you're 50%-50% on Ollie—50% confident that he's O, and 50% confident that he's A. Adam and Ollie undergo the same, unerring test for sameness

[11]Proof: Let $Pr(P) = x$, and let $Pr(Q) = .5$. By the Ratio Formula, we know that $Pr(P|P \equiv Q) = Pr(P \wedge Q)/Pr((P \wedge Q) \vee (\neg P \wedge \neg Q))$. Since P and Q are probabilistically independent relative to Pr, and $(P \wedge Q)$ and $(\neg P \wedge \neg Q)$ are disjoint, $Pr(P \wedge Q)/Pr((P \wedge Q) \vee (\neg P \wedge \neg Q)) = (Pr(P) \times Pr(Q))/((Pr(P) \times Pr(Q)) + (Pr(\neg P) \times Pr(\neg Q)))$. Substituting x for $Pr(P)$ and .5 for $Pr(Q)$, we get: $(.5x)/(.5x + .5(1 - x)) = x$.

[12]You might say that learning that Adam is the same type as Ollie and Arnie really *should* decrease your credence that Adam is A. The lesson of Kidney Transplant is not that belief doesn't reduce to credence, but that we have the wrong theory of credence—rational credence functions are not probability functions. I don't see this as an objection. My claim is that *belief* can't be understood solely in terms of probability. Perhaps Kidney Transplant reveals an even deeper problem with Epistemic Bayesianism—that rational *credence* can't be understood in terms of probability, either. That may be right. But my central point still stands: we can't reduce all facts about rational opinion to subjective probability.

of blood type as before. The test reveals that Ollie and Adam are the same blood type. What are you now entitled to think? Surely you are still entitled to believe that Adam is A. You've learned nothing that contradicts Adam's testimony; your entitlement to trust him is unscathed. You should also infer, from your new belief that Ollie is the same type as Adam, that Ollie must be blood type A, too. If you know Adam's blood type, and you have no information about Ollie's, learning that Adam and Ollie are the same *tells you* Ollie's blood type.

This case shows that sometimes when you're 50%-50% on some proposition Q, and entitled to believe P, learning $P \equiv Q$ *transmits* rational entitlement from P to Q; it's a way of *extending* your beliefs to include Q. The original version of Kidney Transplant shows that things often go the other way around. In that case, you are 50% confident that Ollie is A, and entitled to believe that Adam is A; but learning that Adam is A just in case Ollie is undermines your entitlement to believe Adam is A. Sometimes when you're 50%-50% on Q, learning $P \equiv Q$ *defeats* your entitlement to believe P.

What's the difference between these cases? The answer lies in the nature of your reasons for being 50%-50% on Q. You may be 50%-50% because you have no relevant information about Q. You are *indifferent*. Or you may be 50%-50% because you have symmetrically conflicting evidence about Q. You are *ambivalent*. If you have no relevant evidence about Q, learning $P \equiv Q$ gives you no new evidence about P, so there's nothing to undermine the rationality of your belief in P. If you have conflicting evidence about Q, you have some evidence *against* Q, and this, together with your new belief that P is true only if Q is, defeats your entitlement to believe P.

When you start with no information about Ollie, you should infer from Adam's testimony, together with your knowledge that Adam and Ollie are the same, that Adam and Ollie are both blood type A. Normally people can be trusted to accurately report their blood types, and absent evidence to the contrary, you're entitled to rely on Adam's testimony. When you have conflicting information about Ollie, learning that Adam and Ollie are the same *gives* you evidence to the contrary. Against your new knowledge that Adam and Ollie are the same blood type, Ollie's testimony that he (Ollie) is blood type O is evidence that Adam is O, too. You are no longer entitled to believe that Adam is A.

The probability calculus sees only that you are 50%-50% about whether Ollie is O. It is blind to your *reasons* for being 50%-50%. It treats no information in the same way as it treats symmetrically conflicting information, indifference in the same way as ambivalence. That's because the case of one for type A is probabilistically just like the case of two for type A against one for O. If the two additional pieces of evidence symmetrically conflict—that is, if they support incompatible hypotheses to the same degree—they will cancel each other out. So the overall balance of your evidence—where the evidence points and how decisively it points there—will be the same in the two cases. The larger, more conflicted body of evidence supports the hypothesis that Adam is type A to the same degree as the smaller, unambiguous body of evidence. But your beliefs aren't the same. So whether it is rational to believe a certain hypothesis is not simply a matter of the probability of that hypothesis on your evidence.[13]

2.3.3 Distinguishing Kidney Transplant from Other Cases

I've argued that acquiring symmetrically conflicting evidence about some hypothesis can defeat your entitlement to believe the hypothesis without probabilistically disconfirming it. Now, there are other cases where symmetrically conflicting evidence does disconfirm, and the Simple Theory can easily account for these. One might wonder what's different about Kidney Transplant. Why aren't the resources that the Epistemic Bayesian uses to explain other cases of conflicting evidence

[13]I said that learning that Ollie and Arnie disagree about their shared blood type doesn't reduce your confidence that Adam is type A. That isn't essential. Maybe your credence that Adam is A drops a little bit when you learn that one of Arnie and Ollie is mistaken. But it *doesn't* sink below the level of belief, which is all that's required for Kidney Transplant to be a counterexample to the Simple Theory. You might suggest that your conflicting evidence about Ollie really should push your credence that Adam is A below belief. But that suggestion has implausible consequences. To see why, notice that Kidney Transplant has the following structure. We have one proposition P (Adam is type A), which you believe, and then you gain symmetric, conflicting evidence about some other proposition Q (Ollie is type O). The suggestion, then, would be that in response to conflicting evidence about Q, you ought to drop your belief in P. In general, that can't be right. Suppose you've driven by a new restaurant, Thai Palace. You see the sign on their window from afar, and from what you can make out, it says they're open until 10. You believe that (P) Thai Palace closes at 10. Later on, two friends are arguing about whether Bobby's, a burger drive-in, closes at 10 or 11. You have conflicting evidence about whether (Q) Bobby's closes at 10. When your friends disagree about Bobby's, should you stop believing that Thai Palace closes at 10? Clearly not. You shouldn't worry about *all* restaurants just because you have conflicting evidence about Bobby's.

available to him when it comes to this case?

Suppose you have a twinge of pain in your ankle, and, being a bit of a hypochondriac, you rush to the hospital, where you have an MRI. A doctor examines the MRI and says you have a fracture. You trust the doctor and conclude that you have a fracture. Later, two more doctors review the MRI. One of the new doctors says you have a fracture; the other says it's a sprain. If you trust all three doctors equally, it seems it would be rational to stop believing you have a fracture when you encounter this conflicting evidence.

The Epistemic Bayesian can easily make sense of *this* case. Although the disagreement is symmetric (it's not as though you trust the doctor who says that you have a sprain more than the other two), the fact that there's disagreement is *itself* a sign that the evidence is murky—that it's not so easy to tell fractures from sprains, that the MRI scan was blurry, or perhaps that your injury is a borderline case. Learning that two other doctors disagree can undermine your trust in any doctor who bases his judgment on this evidence, including the first doctor you consulted. This reduces your confidence that you have a fracture. And a reduction in confidence is exactly what the proponent of the Simple Theory needs to predict that you lose your belief.

Kidney Transplant is in *some* ways analogous to this case. Adam is like the first doctor—you start out believing that Adam is type A on the basis of his testimony. One might be tempted to extend the analogy further still, alleging that *against your knowledge that Adam, Ollie, and Arnie are the same blood type*, Ollie and Arnie are like the two disagreeing doctors—just as the doctors disagree about whether you have a fracture, Ollie and Arnie disagree about whether Adam is type A.

But Arnie and Ollie are not just like the disagreeing doctors. The two new doctors testify about the very same question as the first—the question of whether you have a fracture—and importantly, *base their beliefs on exactly the same evidence* as the first doctor. The best explanation of why the doctors disagree is that the evidence is hard to assess—perhaps the MRI scan doesn't clearly tell one way or the other. If disagreement between the two new doctors casts doubt on the quality of their evidence, it also casts doubt on the quality of the *first* doctor's evidence. That's why the fact that they disagree undermines your trust in the first. Things are different in Kidney Transplant. The best explanation of why Ollie and

Arnie disagree is not that blood tests are systematically unreliable, but something much more nebulous—that sometimes people make mistakes about their medical history. But of course you knew *that* all along. Disagreement between Ollie and Arnie about their shared blood type need not cast doubt on your evidence about Adam's blood type, and so learning that Ollie and Arnie are the same type should not undermine your trust in Adam.

Formally speaking, the difference between the doctor case and Kidney Transplant is that *probabilistic independence* is violated in the former, but not in the latter. Whether the first doctor is right about your injury clearly isn't probabilistically independent of what the second two doctors say about it. When you learn that the doctors disagree, you are rationally less confident that the first doctor is right than you would have been otherwise. But Kidney Transplant is different. Whether Adam is right about his blood type *is* probabilistically independent of what Ollie and Arnie say about their own blood types.

Conflicting evidence about a certain hypothesis can undermine one's confidence in it even when the conflict is symmetric, as when neither of the disagreeing doctors is any more trustworthy than the other. A natural explanation of this phenomenon is that disagreement is often a sign that the evidence is equivocal or hard to assess—in one way or another, that it's not to be trusted. That explanation may be satisfactory in the doctor case. But it simply doesn't apply to cases like Kidney Transplant—that there is disagreement about Ollie is not a sign that the evidence about Adam can't be relied on. That the Simple Theory predicts a difference in the undermining effect of conflicting evidence in these two cases shows that it is not an adequate model of conflicting evidence. And if that's so, it's not an adequate account of rational belief, either.

In the next two sections, I consider two more sophisticated accounts of rational belief on behalf of the Epistemic Bayesian. The first theory, the Stability Theory, says that whether you're entitled to believe some proposition P doesn't just depend on how confident you are in P—it also depends on the *stability* of your credence in the face of changing information. The second theory, the Imprecise Theory, retains the idea that high confidence is sufficient for rational belief, but understands confidence in a new way, replacing precise probabilities with sets of probabilities. Neither succeeds.

2.4 The Stability Theory

The *balance* of a body of evidence with respect to a particular proposition P is, roughly speaking, a matter of how decisively the evidence supports P. We've seen that whether you're entitled to believe P isn't entirely determined by the balance of your evidence with respect to P; that is, whether you're entitled to believe P isn't entirely determined by the probability of P on your evidence. In some cases, acquiring conflicting evidence with respect to a proposition P can defeat your belief in P without changing the balance of your evidence—without making it any less probable that P is true.

Perhaps rational belief also depends on what some have called evidential *weight*.[14] The weight of a body of evidence is the gross amount of information available. Two bodies of evidence can have the same balance but different weights. Sam is 80% confident that Serena Williams is going to win the match because he's seen her beat this opponent twice before. John is 80% confident because he knows a great deal about Williams and her opponent—he knows that Williams has won 70% of their 24 contests, that Williams is even more likely to win on hotter days (and today is a scorcher), that her opponent has a mild wrist injury, and so forth. John's and Sam's evidence may well have the same balance, but they don't have the same weight. John has more information, a more substantial basis on which to rest his 80% confidence that Serena will prevail.

Evidential weight is reflected in the *stability* of credences in the face of changing information. John and Sam will not react the same way to Williams losing the first game of the match—Sam's credence will drop quite a bit, while John's won't change much at all. Maybe we should apply this idea to rational belief: although high credence is necessary for belief, it isn't sufficient. One's high credence must also be sufficiently *stable* in response to new information. On this view, it isn't rational for Sam to believe that Williams will win since watching just the first game would make him reconsider; it is rational for John to believe, however, since seeing Williams lose the first game wouldn't worry him much.

This view is not without precedent. Hannes Leitgeb mounts perhaps the most

[14] I borrow the terms 'balance' and 'weight' from Joyce (2005), who attributes the term 'weight' to Keynes (1921).

thorough contemporary defense of the view. He attributes the idea that stability is necessary for rational belief to Hume's *Treatise*.[15] Leitgeb cites Hume scholar Louis Loeb, who argues that stability is *the* distinctive property of Humean belief. For Loeb, Humean beliefs are not 'lively ideas', as many Hume interpreters have thought, but *steady dispositions*, of which occurrent liveliness is just one of many manifestations.[16]

The rough idea behind any version of the Stability Theory is that if it's too easy to change your mind, you don't count as believing rationally. Of course, it would be too much to demand that beliefs be *indefeasible*. It is our lot as reasoners with limited information that we form beliefs that are defeasible, subject to revision in the face of new, better information. But there are limits to when, and how, such reconsideration can occur, according to the Stability Theory. What are those limits? Which propositions must your high credence be stable with respect to? At the very least, it seems, the propositions you think you aren't too unlikely to learn. Your credence shouldn't count as sufficiently stable if it would drop below the threshold of belief upon learning something *consistent with everything you believe*.[17][18]

The Epistemic Bayesian assumes that rational agents update by conditionalization. If you're rational, your posterior credence in P, upon learning Q, is equal to your prior credence in P conditional on Q. To say that your credence in P doesn't drop below the threshold for belief conditional on some proposition Q is to say that your credence in P conditional on Q exceeds t. We've said that not just any Q counts—stability with respect to all Q's is too high a standard. The Q's are limited to those consistent with everything you believe. Your credence in P is stably high just if, for some threshold t, your credence in P exceeds t, and for any Q consistent with your beliefs, your credence in P conditional on Q exceeds t.[19] Formally:

[15] See Leigeb (2013), Leitgeb (2014), and Leitgeb (2017).

[16] See Loeb (2001), Loeb (2002), and Loeb (2010).

[17] One might worry that the Stability Theory is circular. It is. But it is not intended to be an *analysis* of belief, so much as a constraint—whatever belief is, it had better be stable in the sense discussed here. If analysis isn't the goal, circularity needn't be worrying.

[18] This is roughly the proposal offered by Leitgeb (2017) in his *Humean Thesis Explicated*, which goes as follows. Where *Bel* is a rational subject's set of believed propositions, and P is his subjective probability measure: For all X, *Bel*(X) iff for all Y, if $\neg$*Bel*($\neg Y$), and $P(Y) > 0$, then $P(X|Y) > t$, where t is determined by context.

[19] Again I'm assuming $.5 \leq t < 1$.

The Stability Theory: There is some threshold t such that you rationally believes P if and only if, for all propositions Q consistent with your beliefs, $Pr(P|Q) > t$.

2.4.1 Against the Stability Theory

The stability requirement in the Stability Theory is a probabilistic analogue of a principle from the AGM theory of belief revision known as *Rational Monotonicity*:

Rational Monotonicity. If you believe P, then, for any proposition Q consistent with your beliefs, you still believe P upon learning Q.[20]

Rational Monotonicity says that if you believe P, and learning Q would *defeat* your belief in P, you should believe that Q is false. If you believe P, you should believe that you won't encounter any defeaters for that belief. That's not to say that you should think your belief is *indefeasible*. There are things you *could* learn that would make you give up your belief. Rational Monotonicity requires you to believe that you *won't actually* learn any of these things. Your beliefs can be defeated by new information. But if that happens, you shouldn't have seen it coming.

This thought has some intuitive appeal. Suppose I believe that I don't have strep throat because the nurse informed me that my throat culture came back negative. Now, if I learned that throat cultures are unreliable, I would surely renounce that belief. But then it seems I can't *now* leave open this possibility. I can't leave open the possibility that the sole basis for my belief is unreliable; I must believe, at least implicitly, that throat cultures are reliable.

The trouble is that Kidney Transplant is *also* a counterexample to Rational Monotonicity. At Stage Two, you believe that Adam is A. And you leave open that Adam and Ollie are the same (because you leave open that Arnie is right and Ollie is type A). Nevertheless, learning that Adam and Ollie are the same defeats your belief that Adam is type A: you should stop believing that Adam is A upon learning that he and Ollie are the same type.[21]

[20] See, for example, Alchourrón et al (1985).

[21] There are other counterexamples to Rational Monotonicity, which also suggest that the principle is too strong. Here is one adapted from an example due to Ned Hall (1999). It is mid-July in Boston.

But we must tread carefully here. We haven't yet shown that Kidney Transplant is a counterexample to the Stability Theory. Rational Monotonicity says that if you believe that Adam is A at Stage Two, you still believe that Adam is A upon learning that Adam and Ollie are the same. The Stability Theory says that if you believe that Adam is A, then you are still *sufficiently confident* that Adam is A upon learning that Adam and Ollie are the same: your credence that Adam is A doesn't drop below the threshold. But we haven't shown *this* claim to be false. Indeed, we've proven that your credence that Adam is A does not decrease when you learn that Adam and Ollie are the same. If your confidence that Adam is A exceeds the threshold at Stage Two, it exceeds the threshold at Stage Three, too.

But if that's right, we seem to be saddled with a different problem. If you're sufficiently confident that Adam is A at Stage Three—if your confidence exceeds the threshold at Stage Three—won't we wrongly predict that you continue to *believe* that Adam is A when you learn that Adam and Ollie are the same? And isn't that the exact result that we've been trying to avoid? Belief at Stage Two, according to the Stability Theory, requires stably high credence at Stage Two. And stably high credence at Stage Two, in turn, requires high credence at Stage Three. In particular, it requires sufficiently high confidence—confidence that exceeds the threshold for belief. It appears that the Stability Theory is no better off than the Simple Theory. It too predicts that if you believe that Adam is A at Stage Two, you still believe it at Stage Three.

This isn't quite right. Leitgeb (2017) maintains that the threshold for belief *varies* according to the 'context' of the reasoner, which includes his probability function. As the subject gathers more information, the threshold may change. To

I believe that there will be at least one sunny day in the next 12 days, and, it seems, I'm reasonable in doing so. But I don't believe that it will be sunny tomorrow, or the next day, or the day after that. (I haven't checked the current weather, but I do know to expect a good deal of sun during the month of July.) In other words, I leave open that it won't be sunny for the next two days, or for the next three days, and so forth. But now take a number towards the top of what I reasonably leave open—say, 10 days of no sun. If there have been 10 days of no sun, should I believe that there will be sun in the next two days? Plausibly, I shouldn't. But if that's right, we have a counterexample to Rational Monotonicity. I believe that it will be sunny once in the next (say) 12 days, and leave open that there will be no sun for the next (say) 10 days. But were I to learn that there will be be no sun for 10 days, I would *stop* believing that there would be sun within 12 days. See Goodman and Salow (forthcoming) for another counterexample to Rational Monotoncity involving coin flips. See Salow (ms) for much more detailed discussion of Rational Monotonicity.

fix ideas, let's suppose that, at Stage Two of Kidney Transplant, your credence that Adam is A is .91, and $t = .90$. The Stability Theory says that you believe that Adam is A. We've shown that, at Stage Two, $Pr(Ad\,A) = Pr(Ad\,A|Ad\,Ol\,same)$: at Stage Two, your unconditional credence that Adam is A is equal to your credence that Adam is A *conditional on Adam and Ollie being the same*. Since you update by Conditionalization, it follows that your credence that Adam is A upon *learning* that Adam and Ollie are the same, at Stage Three, is equal to your Stage Two credence that Adam is A. But even though your credence that Adam is A remains above .91, it doesn't follow that you continue to believe. The threshold may *increase* between Stage Two and Stage Three.[22]

But if Leitgeb pursues this reply, he runs into a different problem. Suppose that, at Stage One of Kidney Transplant, your friend Cosmo tells you he had the flu shot two years ago. You believe him; you have no reason to think he's lying. Of course you're not certain he had the flu shot two years ago; you're (say) 91% confident. At Stage One and Stage Two of Kidney Transplant, your high credence that he's had the flu shot counts as a rational belief, we'll suppose. But at Stage Three—when you learn that Adam and Ollie are the same—it no longer does. The threshold must increase to at least .91 in order to predict that you stop believing that Adam is A at Stage Three. (We stipulated that you are 91% confident that Adam is A at Stage Three.) As a result, you lose your belief that Cosmo had the flu shot two years ago. But that can't be right. What you've learned about Adam's blood type has nothing to do with your belief that Cosmo had the flu shot.[23]

So the Stability Theory can't be right as it stands. But perhaps the problem resides not in the informal idea that rational beliefs are resilient to changing evi-

[22]In addition to the objection I raise in the main text, there's something unsatisfying about this reply. The main motivation for Leitgeb's Stability Theory is the (allegedly) Humean thought that *beliefs* are stable in the face of new information. But if Leitgeb allows the threshold to change—in particular, to increase—as subjects gather more information, Leitgeb has made it rather easy to lose belief. It's far from clear that the resulting theory respects its Humean foundations.

[23]The more general problem is this. To predict that you believe that Adam is at Stage Two, but not at Stage Three, the defender of the Stability Theory must say that the threshold increases between Stage Two and Stage Three: at Stage Two, it must be below your credence in the proposition that Adam is A, and at Stage Three, it must be above it. Take any proposition P at Stage Two such that (i) you believe P at Stage Two, and (ii) your credence in P doesn't exceed your credence that Adam is A. The Stability Theory will predict that if you stop believing that Adam is A at Stage Three, then you also stop believing P.

dence, but in the particular way we've chosen to implement this idea. Could it be that with a few tweaks to the definition of stability, we can set things right?

I don't think we should hold out hope. The informal idea behind the Stability Theory is that weightier bodies of information can justify belief where lighter bodies do not, even when both bodies of information have the same balance. Kidney Transplant seems to pose a direct challenge to this thought. You're entitled to believe that Adam is A at Stage One, and no longer entitled to believe at Stage Three. But the second body of evidence, the one that doesn't justify belief, is weightier than the first, the one that does justify belief. At Stage One, you have one piece of evidence—Adam's testimony—and you're entitled to believe. At Stage Three, you have *three* pieces of evidence—Adam's testimony, as well as Ollie's testimony, and Arnie's—but you aren't entitled to believe.

Defenders of the Stability Theory point to a real and important phenomenon. The weight of evidence matters. It matters to how attitudes change in response to new information—credences based on weightier evidence *are* more resilient. The mistake is to think that *rational belief* requires weighty evidence and resilient credences. We've seen that more evidence can *undermine* rational belief (even when it leaves the balance intact) if it introduces conflict where previously there was none. At both Stage Two and Stage Three of Kidney Transplant, your evidence supports, say, 90% confidence in the proposition that Adam is type A. But at Stage Two, it does so *uniformly*—your evidence consists of Adam's testimony, nothing more—whereas at Stage Three it does so *ambivalently*—Adam's testimony and Arnie's testimony support type A but Ollie's supports type O. When additional weight brings more ambivalence, it can make believing irrational.

2.5 The Imprecise Theory

So far we've assumed that credal states are best modeled by a single probability function, a function that assigns to each proposition a unique, precise credence. For many Epistemic Bayesians, this is a crude idealization, demanding far more precision than is usually warranted by our messy evidence. These *Imprecise Epistemic Bayesians* propose to model agents not with one probability function, but many—a *set* of credence functions; a kind of credal committee.[24] Each member of your credal committee is a perfect Bayesian agent, assigning a single, precise credence to every proposition. But not you. Your attitude may be 'spread out' over a range of values.

Take the proposition that you'll get sick in the next 6 months. You probably don't have a precise credence in this claim, the Imprecise Bayesian says. Instead, you're imprecise—say, 60%-to-90% confident that you'll catch something within six months. We model you with a set of probability functions: for every number in the interval [.6, .9], some probability function—some member of your credal committee—assigns that credence to the proposition that you will get sick in the next six months.[25]

Sets of probability functions reflect degrees of evidential *specificity*. As Joyce (2005) points out, there are different ways for the evidence to be unspecific. It can be *incomplete* in the sense that it doesn't discriminate one hypothesis from alternative hypotheses, and it can be *ambiguous*, in the sense that it is subject to different interpretations. (Of course both incompleteness and ambiguity are matters of degree.) When your credence in a certain proposition P is imprecise, you have no settled opinion about it. The members of your credal committee represent various, more settled views on P that you might take, corresponding to different interpretations of the evidence (in the case of evidential ambiguity) or different ways of 'filling in the gaps' (in the case of incompleteness).

Perhaps rational belief is a matter of evidential balance *and* evidential specificity. Although your evidence needn't be fully specific with respect to some proposition P to support belief in that proposition, it must be specific *enough* that all members of your credal committee are sufficiently confident. It's rational for you to believe P just in case *every* member of your credal committee assigns to P a probability above the threshold:

The Imprecise Theory. There is some threshold t such that it's rational for

[24] For classic discussions of imprecise credences, see Kyburg (1983) and Levi (1974). See Joyce (2005), Joyce (2011), Moss (2014), Schoenfield (2012), and Sturgeon (2008) for contemporary defenses.

[25] Plausibly, the edges of these ranges are vague. I ignore this complication, since nothing I say turns on it.

you to believe P only if every member of your credal committee exceeds t in P.

Here's how the defender of the Imprecise Theory would explain why you should stop believing that Adam is A in Kidney Transplant. At Stage One and Stage Two, all members of your credal committee are sufficiently confident that Adam is A. When you learn that Adam is the same type as Ollie (and Arnie), some members of your credal committee decrease their confidence that Adam is A. That's why you stop believing.

2.5.1 Modeling Kidney Transplant on the Imprecise Theory

For simplicity we'll pretend that there are five members of your credal committee, Pr_1, Pr_2, Pr_3, Pr_4, and Pr_5. Here's your credal committee at Stage One of Kidney Transplant:

$$Pr_1(\text{Adam A}) = Pr_1(\text{Ollie A}) = Pr_1(\text{Arnie A}) = .90$$
$$Pr_2(\text{Adam A}) = .95;\ Pr_2(\text{Ollie O}) = .95;\ Pr_2(\text{Arnie A}) = .95$$
$$Pr_3(\text{Adam A}) = .90;\ Pr_3(\text{Ollie A}) = .95;\ Pr_3(\text{Arnie A}) = .90$$
$$Pr_4(\text{Adam A}) = .90;\ Pr_4(\text{Ollie A}) = .90;\ Pr_4(\text{Arnie A}) = .95$$
$$Pr_5(\text{Adam A}) = .95;\ Pr_5(\text{Ollie A}) = .90;\ Pr_5(\text{Arnie A}) = .90$$
$$\mathcal{P}(\text{Ollie A}) = \mathcal{P}(\text{Arnie A}) = \mathcal{P}(\text{Adam A}) = [.90, .95]$$

Figure 1. Your credal committee at Stage One

Let's take them in turn. Pr_1 is 90% confident that Adam is A, that Ollie is O, and that Arnie is A at Stage One. Pr_2 is a bit more trusting. He is 95% confident that Adam is A, that Ollie is O, and that Arnie is A. Both Pr_1 and Pr_2 trust Adam, Ollie, and Arnie equally. That's not true of the others. Pr_3 trusts Ollie more than Arnie and Adam—he is 95% confident that Ollie is O, but only 90% confident that Adam is A, and 90% confident that Arnie is A. Pr_4 trusts Arnie more than Adam and Ollie—he is 95% confident that Arnie is A, but only 90% confident that Ollie is O, and 90% confident that Adam is A. Finally, Pr_5 trusts Adam more than Arnie and Ollie—he is 95% confident that Adam is A, but only 90% confident that Arnie is A, and 90% confident that Ollie is O.

With this credal committee, the Imprecise Theory can predict that it's rational to believe that Adam is A at Stage One and Stage Two, but not rational to believe this at Stage Three.

Let me explain. Concentrate on Pr_3, who trusts Ollie more than Arnie (and Adam). When Pr_3 learns that Arnie and Ollie are the same type, at Stage Two, he is more confident that they're type O than that they're type A: he thinks Arnie is more likely to be mistaken than Ollie. (Pr_3's credence that Arnie and Ollie are type A at Stage Two is approximately .32.) This means that, upon learning that *Adam* is the same type as Ollie (and Arnie), at Stage Three, Pr_3's credence that Adam is A goes down. Because he's more confident than not that Ollie is type O, the fact that Adam is the same type as Ollie is some evidence that Adam is type O, too. Conditional on Adam sharing his blood type with Ollie (and Arnie), Pr_3 is only .81 confident that Adam is type A.[26]

Let's take a step back. The Simple Theory couldn't make sense of Kidney Transplant because, on the precise picture, if you are no more confident in Ollie than Arnie, and *vice versa*, then when you learn they're the same, you will be 50% confident that Arnie is right and they're both A, and 50% confident that Ollie is right and they're both O. As a result, when you learn that Adam is the same type as Ollie, you effectively gain no new evidence about Adam. Your credence that Adam is type A doesn't budge.

Things are different for the Imprecise Theory. You are still no more confident in Ollie than Arnie, and *vice versa*. When Ollie tells you he is type O, you are 90%-95% confident that he is O; when Arnie tells you he (Arnie) is A, you are 90%-95% confident of that, too. But—and this is the important part—that doesn't mean you are 50%-50% when you learn that Ollie and Arnie are the same type. *Some* members of your committee—those who trust Arnie and Ollie equally—are 50%-50%. But others aren't. Pr_3, who trusts Ollie more than Arnie, is 68% confident that Arnie and Ollie are both A. Pr_3 *does* acquire evidence that Adam is O when he learns that Adam is the same type as Ollie. And *your* beliefs are sensitive to his evidence, and his confidence. If his confidence drops below the

[26] Calculation. $Pr_3(Ad\,A|Ad\,Ol\,same) = Pr_3(Ad\,A \wedge Ol\,A)/Pr_3((Ad\,A \wedge Ol\,A) \vee (Ad\,O \wedge Ol\,O)) = (Pr_3(Ad\,A) \times Pr_3(Ol A))/((Pr_3(Ad\,A) \times Pr_3(Ol\,A)) + (Pr_3(Ad\,O) \times Pr_3(Ol\,O))) = (.9 \times .32)/((.9 \times .32) + (.1 \times .68)) \approx .81$.

threshold, you stop believing.

2.5.2 Against The Imprecise Theory

The Imprecise Theory fares better than the Simple Theory or the Stability Theory. But it isn't enough. Though the Imprecise Theory delivers the right verdicts in some versions of Kidney Transplant, the theory founders when we apply it to others. There's a clear explanation for the Imprecise Theory's shortcomings: it is predicated on assumptions about Kidney Transplant that are inessential to the case, ones we can readily discard without changing our intuitions about it.

Here's what I mean. To predict that you are no longer entitled to believe that Adam is A at Stage Three, we had to assume that you aren't sure how reliable each of Adam, Arnie, and Ollie is. We also had to suppose that you aren't sure how their respective degrees of reliability *compare*—in particular, some members of your committee had to trust Ollie more than Arnie. Without this last assumption, the Imprecise Theory can't account for the undermining effect of conflicting evidence in Kidney Transplant; it can't explain why you should stop believing that Adam is A at Stage Three. This is a serious limitation of the view. Whether you are entitled to believe that Adam is A at Stage Three doesn't depend on whether you leave open that Ollie is more reliable than Arnie. Even when you are *sure* that Arnie is marginally more reliable than Ollie, learning that Adam is the same type as Ollie (and Arnie) defeats your belief.

Consider a modified version of Kidney Transplant. Adam, Arnie, and Ollie each take a test to determine their blood types. The test is reliable, but known to be a *slightly* more reliable test for type A than it is for type O. Just over half of all errors have been misclassifications of type O individuals as type A, and you know this.

Adam, Ollie, and Arnie report back on this first test. Adam says he is A; Ollie says he is O; Arnie says he is A. You're entitled to believe each of them. They take another test—the one we mentioned earlier—that determines whether two people have the same blood type. (Remember: this test is always right.) You are surprised to learn first that Arnie and Ollie are the same, and just after that Adam and Ollie are the same.

It is still our firm intuition, it seems to me, that you are entitled to believe that Adam is A at Stage One and Stage Two, and that you are *not* entitled to believe this upon learning, at Stage Three, that Adam is the same type as Ollie. To be sure, you should be more confident that Adam is A at Stage Three of this version than you are at Stage Three of the original. But this additional confidence does not save your entitlement to believe that Adam is A. It does not compensate for what is lost when you learn that Ollie's testimony is in conflict with Adam's.

But if the Imprecise Theory were right, the fact that you trust Arnie marginally more than Ollie should make a world of difference. If you are sure that Arnie is marginally more reliable than Ollie, then all of your committee members will be just over 50% confident that Ollie is blood type A when you learn that Arnie and Ollie are the same type. And if that's right, learning that Adam is the same type as Ollie can only give you more evidence that Adam is type A. Every member of your credal committee will *increase* his confidence that Adam is A, and the Imprecise Theory will say that it's rational to keep believing.

Conflicting evidence need not be a kind of evidential imprecision. Or, at the very least, it need not be the kind of evidential imprecision it would have to be if the Imprecise Theory were right. The Imprecise Theory says that acquiring conflicting evidence about whether Adam is A defeats your belief that Adam is A only when your evidence doesn't specify how to *weigh* Arnie's testimony against Ollie's—or, more precisely, only when your evidence *leaves open* that Ollie's testimony counts for more than Arnie's.[27] But, intuitively, that isn't essential. Even if your evidence *determines* that Arnie's testimony confirms the hypothesis that Adam is A more than Ollie's disconfirms it—that is, even if you are *sure* that Arnie's testimony is marginally more reliable that Ollie's—learning that Ollie, someone you antecedently trusted, disagrees with Adam can undermine your entitlement to trust Adam.

[27] Assuming, as always, that the proposition you believe and the proposition for which you have conflicting evidence are probabilistically independent.

2.6 Conclusion

I've argued that conflicting evidence poses a serious challenge for probabilistic theories of rational belief. Conflicting evidence is the sort of evidence you have when two sources of information contradict each other, such as when doctors offer conflicting diagnoses, or when film critics you trust write discordant reviews. Under the right circumstances, conflicting evidence can defeat your entitlement to believe a certain hypothesis without making it any less probable. I presented a case—Kidney Transplant—that illustrates this phenomenon. Kidney Transplant is a counterexample to the simplest probabilistic theory of rational belief—the Simple Theory—according to which rational belief just is rational credence above a certain threshold.

I also considered two more sophisticated probabilistic theories of belief—the Stability Theory and the Imprecise Theory, respectively. We've seen that these theories can't be right, either. The Stability Theory says that high confidence is not enough for rational belief; your belief must also be sufficiently stable in response to new information. But our counterexample to the Simple Theory—Kidney Transplant—is also a counterexample to the idea that rational belief requires stable belief. The Imprecise Theory retains the idea that rational high confidence is sufficient for rational belief, but understands confidence in a different way, replacing single probabilities with sets of probabilities. The Imprecise Theory fails because it wrongly assimilates conflicting evidence to evidential imprecision.

In the wake of these three failed attempts to reduce belief to subjective probability, what are we to think? It would be hasty to conclude that no adequate probabilistic theory of rational opinion will ever be devised. We haven't exhausted the possibilities, and perhaps some, more sophisticated reduction of belief to credence will prove resistant to my objections. Even if we can't rule this out, I think we can say, with confidence, that Epistemic Bayesians have their work cut out for them. If there is a way to reduce belief to subjective probability, the reduction is much messier, much more *global*, than we've been led to believe. Whether it's rational for you to believe P isn't determined by your subjective probability in P. And it isn't determined by your subjective probability in P conditional on various propositions you leave open. If there's a true, purely probabilistic theory of rational opinion, I doubt that it will look much like the theories we have now.

Chapter 3

Can We Decide to Believe?

It is widely held that we cannot believe at will. It is also widely held that this is not just a contingent psychological disability, but a metaphysical necessary limitation grounded in the nature of belief. In this way, belief is said to contrast with blushing. As it happens, blushing is not amenable prodding of the will, but it might have been. We can imagine someone who has a special ability to blush on command. But we cannot imagine someone who forms a belief on command. A belief formed at will would not be a belief at all. Or so it is commonly thought.

Bernard Williams is the source of most contemporary discussion of this issue. His reasons for thinking that it is impossible to believe at will are summarized in the following passage:

> One reason is connected with the characteristic feature of beliefs that they aim at truth. If I could acquire a belief at will, I could acquire it whether it was true or not; moreover I would know that I could acquire it whether it was true or not. If in full consciousness I could will to acquire a 'belief' irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e. as something purporting to represent reality. At the very least, there must be a restriction on what is the case after the event; since I could not then, in full consciousness, regard this as a belief of mine, i.e. something I take to be true, and also know that I acquired it at will. With regard to no belief could I know—or, if all this is to be done in full consciousness even

suspect—that I had acquired it at will. But if I can acquire beliefs at will, I must know that I am able to do this; and could I know that I was capable of this feat, if with regard to every feat of this kind which I had performed I necessarily had to believe it had not taken place? (Williams 1970, 148)

It seems to me that there are two importantly different arguments contained in this passage. The first argument, which I will call the *aim-of-belief argument*, focuses on the act of *forming* a belief at will. To form a belief intentionally is to form a belief with a certain *aim*—the aim of believing the truth. To try to form a belief is to try to form a true belief, and you can't do *that* by forming a belief at will. The second argument—I will call it the *retrospective argument*—focuses on the state of believing, rather than the act of forming a belief. If you were to form a belief at will, it says, you would know something about the belief that is inconsistent with holding the belief. That's why it isn't possible.

I argue that neither argument succeeds in showing that it is impossible to believe at will. The retrospective argument allows that you can decide to believe in acknowledged permissive cases, those in which you believe that your evidence permits believing and permits not believing. The aim-of-belief argument allows that we can change our doxastic states at will when they are imprecise.

3.1 The Retrospective Argument

The retrospective argument is found in the second half of the passage quoted above. Here it is again:

> With regard to no belief could I know—or, if all this is to be done in full consciousness, even suspect—that I had acquired it at will. But if I can acquire beliefs at will, I must know that I am able to do this; and could I know that I was capable of this feat, if with regard to every feat of this kind which I had performed I necessarily had to believe it had not taken place? (Williams 1970, 148).

The argument appears to have two premises:[1]

[1] I borrow this reconstruction from Setiya (2008).

(1) If I am able to believe at will, I know that I am able to believe at will.

(2) I cannot at once believe P and know that my belief in P was formed at will.

Williams seems to think it a consequence of (2) that I cannot know that I am able to form a belief in P at will. But if I cannot know that I am able to form a belief in P at will, it follows from (1) that I am not able to do so.

There are several flaws in the argument. For starters, the argument is invalid—at last as stated. To see this, suppose we grant premise (2)—that I cannot at once believe P and know that my belief in P was formed at will. Does it follow that I cannot know that I am able to believe at will? As Jonathan Bennett (1990) has shown, it does not. Bennett imagines a community of creatures called 'Credamites'. When a Credamite forms a belief at will, he immediately forgets that he has done so. Each act of forming a belief at will is accompanied by local amnesia, in which he forgets how his belief is formed. So, he never knows, of any particular belief, that it was formed at will—the Credamites satisfy (2). But they might nevertheless know that they have the *capacity* to form beliefs at will. It is not a consequence of (2), Bennett concludes, that one cannot know that one is able to form beliefs at will.

And that's not the only problem with the argument. Premise (2) is itself dubious. It says that one cannot believe P while knowing that one's belief was formed at will. But why think that's true? Couldn't I know that a belief of mine formed at will, yet persist in holding the belief if I believed that it was *now* supported by sufficient evidence? (This objection is due to Barbara Winters (1979).)

Despite these objections, it has seemed to many philosophers that Williams was onto something. Kieran Setiya (2008) proposes a refurbished version of Williams's retrospective argument. Let's examine the argument in detail.

Begin with Setiya's definition of believing at will. He proposes the following:

> To believe at will is to form the belief that P by intentional action, believing throughout that, if one were to form the belief or to become more confident that P intentionally, one's degree of confidence or belief would not be epistemically justified.

This definition marks two requirements on forming a belief at will. First, the belief must be formed intentionally. Note that intentionally forming a belief is not the same as intentionally *causing* oneself to form the belief. I do not intentionally form the belief that P when I consult a hypnotist, just as I do not intentionally blush by intentionally dropping my pants in public. When I blush upon dropping my pants, my blushing is the product of intentional action, but is not itself intentional. The same goes for belief. When I consult a hypnotist, my forming the belief that P is the product of my intentional action, but is not itself intentional.

What about the second requirement? It says you form a belief in P at will only if you believe that, if you were to form the belief intentionally, your belief would be unjustified. This requirement is meant to capture the idea that believing P at will is done without regard for the truth of P.

Setiya's first step towards argument against the possibility of believing at will in this sense is to replace Williams's (2) with (2*) below:

(2*) It is impossible to believe that P or to be confident that P while believing that this degree of confidence or belief is not epistemically justified.

We objected to Williams's (2) on the grounds that one can know perfectly well that one's belief was formed at will, so long as one takes it now to be supported by sufficient evidence. What one thinks about the belief's *origins* is neither here nor there. But this objection implicitly concedes that what one thinks about one's belief's *present* epistemic status does matter. One cannot at once believe something while believing that one does not have sufficient evidence for this belief. Setiya's (2*) is meant to capture this idea.

Notice that if we accept (2*), as well as Setiya's definition of believing at will, we must also accept that one cannot know that one's belief was formed at will. But what if, as Bennett suggested, one forgets the origins of one's beliefs the moment they are formed? So long as the amnesia kicks in on time, one will never find oneself in the predicament ruled out by (2*). Williams thought that (1) excluded this possibility. But he was wrong. As we said, one can know that one has the capacity to form beliefs at will without knowing, of any particular belief, that it was formed at will.

How, then, are we to rule out Bennett's Credamites? Setiya's suggestion is to appeal to the nature of intentional action. He replaces Williams's (1) with (1*):

(1*) If A is ϕ-ing intentionally, he believes that he is doing so, or else he is doing by performing some other intentional action, which he believes he is doing.

Following Setiya, let's consider a specific attempt to believe at will, as when Alice undertakes the White Queen's challenge to believe that 'she is just one hundred and one [years], five months and a day.' It follows from the definition of believing at will that Alice forms her belief by intentional action.

Now, to form a belief in a certain proposition, Setiya thinks, just is to become more confident in that proposition. That is the argument's third, and final, premise:

(3*) To form a belief in P is to become more confident in P.

This premise tells us that for Alice to form the belief that the White Queen is a hundred and one just is for Alice to become more confident in this proposition. We know from (1*) that Alice is intentionally forming the belief that the White Queen is a hundred and one. Only a sort of conceptual confusion could prevent her from intentionally becoming more confident that the White Queen is a hundred and one. But, as Williams insists, what is done only as a result of conceptual confusion is not done in full consciousness, and so we should set aside the possibility that Alice is confused about what it is to form a belief.

We've established that Alice has intentionally become more confident that the White Queen is a hundred and one. Now, becoming more confident in some proposition is plausibly a *basic* action—it is not the kind of thing you do *by* doing something else. If that's right, it follows from (1*) that Alice believes she is intentionally becoming more confident that the White Queen is a hundred and one. But, in general, if you are becoming more confident in P, you have become more confident in P. (Linguistically, 'becoming more confident' is like 'walking' or 'singing': If you are walking, you have walked; if you are singing, you have sung.) So, again barring conceptual confusion, if Alice believes she is intentionally becoming more confident that the White Queen is a hundred and one, she believes she has intentionally become more confident that the White Queen is a hundred and one.

But now remember the second requirement on believing at will. It says that you believe throughout that if you were to form the belief or become more confident that P intentionally, your belief would be unjustified. It looks like Alice finds herself in the following predicament. She believes she has intentionally become more confident that the White Queen is a hundred and one, and she also believes that if she were to intentionally become more confident in this proposition, her confidence would not be justified. Absent confusion or inattention, she can deduce that she is confident that the White Queen is a hundred and one and that her confidence is unjustified. But (2*) says that this state is impossible. That is why Alice cannot, at will, meet the White Queen's demands.[2]

Though I'm sympathetic to Setiya's premises, I don't think he has shown we can't believe at will. That's because I don't think Setiya has provided an adequate account of what it is to believe at will. I won't take issue with the first requirement—that a belief formed at will is a belief formed intentionally. That part is indispensable, a minimal condition on what counts as believing at will. But the second requirement must go.

Remember that the second requirement is meant capture the idea that a belief formed at will is a belief formed without regard the truth of the belief. Setiya is surely right that *one* way to form a belief without regard for the truth is to do so while knowing that if you were to form the belief, it would be irrational. But it's not the only way. Suppose you believe that your evidence is *permissive* with respect to some proposition—say, the proposition that God exists. You believe that it is rational to believe that God exists and you believe that it is rational to believe that God doesn't exist. Unsure what to do, you flip a coin, giving yourself the following instructions: believe in God if the coin lands heads; otherwise, believe that God doesn't exist. The coin lands heads, and on that basis, you form the intention to believe that God exists, and thereby come to believe it. Surely if anything counts as forming a belief at will, forming a belief by flipping a coin does. But not according to Setiya's definition. As far as it's concerned, forming a belief by flipping a coin in a believed permissive case is not an instance of believing at will.[3]

[2]This argument assumes that if Alice intentionally forms the belief that the White Queen is a hundred and one, she is at some point intentionally *forming* the belief that the White Queen is a hundred and one. But what about instantaneous intentional belief-formation? Could Alice form her belief by undergoing an intentional change of state that has no duration at all? Setiya allows that this may be possible, but insists that it doesn't matter. If one is performing a non-durative action, he says, one knows that one is doing it. I won't take issue with this claim here.

[3]For defenses of Epistemic Permissivism—the view that there are permissive cases like the one I've described in the main text—see Kelly (2013), Meacham (2013), Schoenfield (2014), Schoenfield (2017), and Titelbaum and Kopec (forthcoming). For criticisms of Epistemic Permissivism, see

But maybe this objection overlooks an important distinction. Epistemologists typically recognize two kinds of epistemic justification: *propositional* justification and *doxastic* justification. You are propositionally justified in believing P just in case you have sufficient reason to believe P, whether you believe it or not (and, if you do believe, whether or not you do so for good reasons).[4] Doxastic justification is more demanding. You are doxastically justified in believing P just in case you have sufficient reason to believe P and you base your belief in P on those reasons. The juror who has overwhelming evidence that Jones is guilty, but bases her belief on the fact that he looks suspicious is propositionally, but not doxastically, justified.

With this distinction in hand, return to the example. You're about to flip a coin to decide whether to believe that God exists: believe if it lands heads, otherwise don't. We said that you believe that your evidence is permissive with respect to the proposition that God exists. You believe that you have sufficient reason to believe that God exists and you have sufficient reason to believe that God doesn't exist. In other words, you believe that you have propositional justification for believing that God exists, and that you have propositional justification for believing that God doesn't exist. And, we'll suppose, you know that flipping the coin won't change this. So you believe that if you were to form your belief on the basis of the coin flip, your belief would be propositionally justified.

But do you also believe that it would be *doxastically* justified? Perhaps not. Doxastically justified beliefs are typically not formed on the basis of coin flips, after all. So maybe Setiya's argument for the impossibility of believing at will eludes my objections if we state his definition of believing at will in terms of doxastic justification, as I've done below:

> To believe at will is to form the belief that P by intentional action, believing throughout that, if one were to form the belief or to become more confident that P intentionally, one's degree of confidence or belief would not be doxastically justified.

Greco and Hedden (forthcoming), Horowitz (2013), Horowitz (forthcoming), Schultheis (forthcoming), and White (2007).

[4]This is the orthodox definition anyway. See, among others, Feldman (2002), Kvanvig (2003), and Pollock and Cruz (1999) for defenses. See John Turri (2010) for objections to the orthodox account, and for a different proposal, which defines propositional justification in terms of doxastic justification, rather than the other way around.

Does this definition count forming a belief in God on the basis of a coin flip as an instance of believing at will? That depends on whether the subject believes that if she were to believe that God exists on the basis of a coin flip, her belief would be doxastically justified. Advocates of causal theories of doxastic justification say that one is doxastically justified in believing some proposition only if one's belief is *causally sustained* by the source of one's propositional justification, not by other factors. They will of course insist that this condition is violated when one knowingly forms a belief on the basis of a coin flip.

But what *they* think about the nature of doxastic justification is beside the point. What matters is what the *subject*—the one who's about to flip the coin—believes. She may reject the causal theory of doxastic justification. It's not the only theory, after all. The *doxastic theory* of doxastic justification, in contrast with the causal theory, says that believing, or rationally believing, that your belief is propositionally justified is sufficient for justifiably holding the belief. The belief need not be causally sustained by the source of propositional justification.

Why would one hold this view? Because there seem to be counterexamples to the causal theory. Suppose you have strong evidence that you're going to win a local lottery; you've been told that it's rigged in your favor, and you regard this as conclusive evidence that you will win. But you're also prone to wishful thinking, and you would believe that you were going to win even if you had no evidence. Are you justified in holding your belief that you will win the lottery? It seems that you are. You rationally believe—perhaps you even *know*—that you have strong evidence that you are going to win. If questioned, you would be prepared to defend your belief by appeal to what is, by hypothesis, sufficient evidence. What more could doxastic justification demand?[5]

Maybe our subject, ready to flip her coin, holds the doxastic theory. She believes that she would be doxastically justified were she to form a belief on the basis of a coin flip in what she takes to be a permissive case. She knows now, and she knows that she will continue to know, that her belief is propositionally justified. And she knows that, if queried, she would be ready to justify her theism, citing arguments and facts which, taken together, provide sufficient reason to believe. To

[5] This is Setiya's own example. See Setiya (2013), page 191. Similar examples have been given by Lehrer (1971) and Frankish (2007).

her mind, that is enough to show that she would be justified in holding the belief that God exists.

Whatever we mean to rule out when we say that believing at will is impossible, forming a belief by flipping a coin in a believed permissive case is surely one of those things. But why, exactly? I've argued that it's not because you expect your belief to be unjustified. Since you know the case is permissive, you know that your belief would be (propositionally and doxastically) justified.

Maybe it has something to do with the fact that, when you decide what to believe by flipping a coin, you know that your reason for forming the belief does not bear on the truth of the belief. This yields the following definition of believing at will.

> To form a believe at will is to form the belief that P by intentional action, believing throughout that one's reason for forming the belief that P or for becoming more confident in P does not bear bear on the truth of P.

Let's try to build an argument for the impossibility of believing at will in this sense. We can reuse Setiya's first premise, repeated below:

(1*) If A is doing something intentionally, he believes that he is doing it, or else he is doing it by performing some other intentional action, which he believes he is doing.

Go back to Alice, who wishes to believe that the White Queen is a hundred and one. We said earlier that if Alice forms the belief intentionally, she also intentionally becomes confident that the White Queen is a hundred and one. (We're assuming that Alice is not confused about the nature of belief; she knows that to form a belief just is to become more confident.) We also said that if she is intentionally *becoming* more confident that the White Queen is a hundred and one, she has intentionally *become* more confident. And (1*) ensures that she knows this. From the definition of believing at will, we know that Alice believes that her reason for forming the belief that the White Queen is a hundred and one does not bear on the truth of the proposition. Putting everything together, we can see that if Alice forms her belief at will, then she believes that she has intentionally become more confident that the White Queen is a hundred and one, and she believes that her reason for doing so does not bear on the truth of this proposition.

Our final step is to say that this state is impossible. But we can no longer use Setiya's (2*) for this purpose. We must replace it with something like this:

(2**) It is impossible to believe P or be confident in P while knowing that your reason for forming the belief that P or for becoming more confident in P does not bear on the truth of P.[6]

But I don't see why we should accept (2**), for reasons mentioned earlier. Alice can know perfectly well that her belief was formed for some reason that doesn't bear on its truth so long as she believes that it is *now* supported by sufficient evidence. What she knows about how the belief was formed is irrelevant.

We might try to repair the argument by replacing (2**) with (2***):

(2***) It is impossible to believe P or be confident in P while knowing that your reason for believing P or for being confident in P does not bear on the truth of P.

The difference between (2**) and (2***) is subtle. While (2**) talks about reasons for forming beliefs and becoming more confident, (2***) talks about reasons for *believing* and reasons for *being confident*. Perhaps, then, (2***) is not subject to the particular objection we raised against (2**).

But even if (2***) were true, it's far from clear why it should apply to Alice. When Alice forms the belief that the White Queen is a hundred and one at will, she forms her belief for some reason that doesn't bear on the truth of the proposition, but it doesn't follow that you *believes* for reasons that don't bear on the truth of the proposition. Forming a belief and believing are two different things—one is an action, the other a state. We can't just assume that one's reasons for performing the action are also one's reasons for being in the state.

[6]This is close to the principle Barbara Winters proposes in place of Williams's original second premise. Her premise says that it is not possible to believe, in full consciousness, both that one believes that P and that one's belief that P is not currently sustained by truth-considerations.

The guiding idea behind the retrospective argument was this. If you were to form a belief at will, you would know something about your belief that is inconsistent with holding it. Believing at will is self-defeating—any attempt to do it would thwart its own end. What is this piece of knowledge that would dissolve the belief the moment it is formed? The only plausible answer is knowledge of the belief's irrationality. You can't maintain your belief while acknowledging that it is irrational, but this is exactly the state you would be in were you to form a belief at will.

I have argued that this last step is unwarranted. It may well be impossible to maintain a belief that you think is irrational, but nothing about the act of forming a belief at will, properly understood, forces you into this predicament. When you form a belief by flipping a coin in what you take to be a permissive case, you are forming it at will, but you don't expect your belief to be irrational.

3.2 The Aim of Belief

We began with a passage from Bernard Williams, observing that it contained not one, but two importantly different arguments for the conclusion that it is impossible to believe at will. I have argued that one of these arguments—the retrospective argument—doesn't work. It doesn't explain why we can't believe at will in acknowledged permissive cases. What about the second argument—the aim-of-belief argument?

When Williams says that belief aims at truth, he means that belief *constitutively* aims at truth. It is part of what a belief *is* that it is a mental state that is directed at the truth. But what, exactly, does it mean to say that belief aims at truth? Different philosophers say different things. According to some, to say that belief aims at truth is simply to say that to believe some proposition p is to regard p as true. But this doesn't distinguish belief from other attitudes, like imagining or supposing. To suppose p is to treat p as true for the sake of argument; to imagine p is also to regard p as true.[7] According to others, to say that belief constitutively aims at truth is to say something about how belief is *regulated*—how it is formed, revised,

[7] Velleman (2000) also makes this point.

and extinguished. Typically, one forms beliefs because one has evidence that they are true, and abandons them when one acquires evidence of their falsity. Belief, as Velleman and Shah (2005) put it, is regulated for truth.

But this characterization of the way in which belief aims at truth will not help us account for the impossibility of believing at will. To say that belief is regulated for truth is to say something about how beliefs are normally or typically formed, not how they *must* be formed. It is surely possible for one's beliefs to be influenced by considerations that don't bear on their truth—wishful thinking is all too common. And so for all we've said, it's also possible for non-alethic considerations to play a role in doxastic deliberation, to borrow a term from Velleman and Shah (2005). For all we've said, it is possible to decide to believe on the basis of a coin flip.

A different suggestion comes from Velleman (2000). He observes that supposing, imagining, and believing p all involve treating p as true—*accepting* p—for a certain purpose. To suppose p, for example, is to often to accept p for the sake of argument. To believe p is also to accept p, Velleman says, but not for the sake of argument—it is to accept p with the aim of accepting p if and only if p is true. To say that belief constitutively aims at truth, on Velleman's view, is to say that a state of accepting p does not count as a belief in p unless the subject accepts p with the aim of believing the truth about p. But as Velleman himself later observes, this can't be right, either. Most instances of belief-formation don't involve any aims or goals at all. I form the belief that the wall is red or that the sun is shining not because I have a certain aim or goal—it just happens automatically when I perceive that the wall is red or that the sun is shining.

Here is what I propose. To say that belief aims at truth—in the sense that is relevant to question of whether we can believe at will, that is—is not to say that one cannot *form* a belief in p unless one is doing so with the aim of believing p if and only if p is true. It is to say that one cannot *intentionally* form a belief in p unless one is doing so with the aim of believing p if and only if p.

This characterization seems right when we consider other aim-directed activities, such as spelling. The activity of spelling 'cemetery' aims at writing (or uttering, as the case may be) the letters of the word 'cemetery' in the order I've just written them in. This does not mean that one cannot spell incorrectly. Nor does it mean that how one spells cannot be influenced by factors that have nothing to do with correct spelling. When we say that correct spelling is a constitutive aim of spelling, we mean at most that one cannot intentionally spell unless one is doing so with the aim of spelling correctly. That is why I cannot try to spell 'cemetery' by scribbling 'FFF' on the page. This would not be an act of trying to spell 'cemetery' correctly, and so it would not be an act of trying to spell at all.[8]

To say that correct spelling is a constitutive aim of spelling, I propose, is to say that one does not count as trying to spell a word unless one is trying to spell it correctly. Similarly, one doesn't count as trying to form a belief in p unless one is trying to form a true belief in p—that is, unless one is doing so with the aim of believing p if and only if p. Perhaps the fact that belief aims at truth in *this* sense explains why I cannot form a belief at will. If I were to decide what to believe on the basis of a coin flip, I would not be trying to form a true belief, and so I wouldn't be trying to form a belief at all.

There are two problems with this argument. The first is that its conclusion is not general enough. It says that we cannot form beliefs at will. But beliefs aren't the only doxastic attitudes we have. We also have credences, levels of confidence, degrees of belief, as they are variously called. And presumably, these are no more responsive to the prodding of the will than our beliefs are. Just as I cannot, just like that, form the belief that Los Angeles is the capital of California, I cannot, just like that, raise my confidence in this proposition. Whatever we mean to rule out when we say that we cannot believe at will, it includes our credences.

The second problem is that the argument is invalid—as it stands, at least. To see why, return to the spelling analogy. Let's grant that spelling constitutively aims at correct spelling. If I'm not trying to spell correctly, I'm not trying to spell at all. Does it follow from this fact that I cannot choose, on the basis of a coin flip, how to spell 'cemetery'? Clearly not. Suppose I've forgotten how 'cemetery' is spelled. Vacillating between 'C-E-M-E-T-E-R-Y' and 'C-E-M-E-T-A-R-Y', I decide to flip a coin, giving myself the following instructions: if the coin lands heads, spell 'cemetery' 'C-E-M-E-T-E-R-Y'; if it lands tails, spell it 'C-E-M-E-T-

[8]One might object that one can intentionally spell incorrectly. Suppose you ask me how to spell my name, but for some reason I don't want you to know how to spell it, so I intentionally spell my name incorrectly. I think we can account for this by positing an ambiguity in the verb 'to spell'. In one sense of the term, I have spelled incorrectly; in another sense, I wasn't really trying to spell my name, since I wasn't trying to spell it correctly.

A-R-Y'.

Is it not possible for me to comply with these instructions? Am I not free to decide, on the basis of how the coin lands, how to spell 'cemetery'? Surely I *can* comply with my instructions; I *am* free to decide how to spell 'cemetery.' Things would be different if I knew how to spell 'cemetery'. When I know that 'cemetery' is spelled 'C-E-M-E-T-E-R-Y', I can't try to spell it any other way. But if I forget, I gain some control over how I spell the word. If the coin lands heads, I can decide to spell 'cemetery' 'C-E-M-E-T-E-R-Y'; if it lands tails, I can decide to spell it 'C-E-M-E-T-A-R-Y'.

The lesson is that, even though spelling aims at correct spelling, I have some discretion over how I spell words when I don't know how to spell them—that is, when I don't know what best satisfies my aim of correct spelling. When 'C-E-M-E-T-E-R-Y' and 'C-E-M-E-T-A-R-Y' seem like equally plausible candidates for how 'cemetery' is spelled, they are, from my perspective, equally good ways to satisfy my aim, and so I can choose to spell the word either way.

The same goes for other aim-directed activities, like singing the lyrics to a song. Suppose I don't know how whether the lyrics to the Monkees' 'I'm a Believer' are 'then I saw her face, now I'm a believer' (the actual lyrics) or the mistaken 'then I saw her face, now I'm gonna leave her.' I've heard the song many times, but just can't tell which is right. Surely I can decide which to sing on the basis of a coin flip.

But if belief is anything like spelling or singing a song, we should expect something similar to be true of belief. We should expect there to be cases in which I don't know what best satisfies my aim of believing the truth. And in those cases, we should expect to find that I have some discretion over what I believe. We should expect that I can decide whether to believe on the basis of a coin flip.[9]

[9]Note that we can't solve this problem by changing the aim—e.g. by saying that one's aim is to not to believe the truth about P, but to believe P only if P is true. If you believe P is true, you can satisfy the aim of believing P only if P is true by believing P and by being agnostic. Why not flip a coin to decide which to do?

3.3 Accuracy and the Aim of Credence

There were two problems with the aim-of-belief argument. It's conclusion is too narrow—it doesn't say anything about raising one's confidence at will, but we want to rule that out, too. And if you are uncertain about what best satisfies your aim, you have some discretion over how you pursue it.

The only way to solve both problems is to refashion the argument in a fine-grained framework in which we have, in addition to beliefs, varying levels of confidence, or credences. Now, credences can't be true or false, so whatever they aim at, it's not truth. A popular idea among contemporary epistemologists is that credences aim at *accuracy*—a graded surrogate for truth. Accuracy is like truth, they say, but it comes in degrees. Credences are more or less accurate, depending on how close they are to the truth. An entire branch of contemporary epistemology, known as *epistemic utility theory*, has grown up to develop this idea in precise terms.

Let me briefly introduce the framework. We said that credences aim at accuracy. But what is accuracy and how do we measure it? We will think of the accuracy of a subject's credence in a certain proposition as a measure of how close it is to the truth about that proposition. Your credence is maximally accurate when you have credence 1 in a truth or 0 in a falsehood, maximally inaccurate when you have credence 1 in a falsehood or 0 in a truth. The greater your credence in P, the more accurate you are with respect to P when P is true, and the less accurate you are with respect to P when P is false.

We can also measure the accuracy of a subject's *credence function*, which takes propositions about which she has an opinion, and returns her credence in those propositions. A credence function c is maximally accurate when it assigns 1 to all the truths and 0 to all the falsehoods. It is minimally accurate when it assigns 1 to all the falsehoods and 0 to all the truths. In general, the higher your credences are in truths, and the lower they are in falsehoods, the more accurate your credence function will be.

A bit more precisely now. Let S be a partition of possible states of the world, and let C_s be a credence function defined over S. We measure accuracy using a scoring rule A, which takes a state of the world s, from the partition S, and a credence function c, from the set of credence functions C_s, and returns a number between 0 and 1. This number represents the accuracy of c is in s.

$$\mathbf{A}: C_s \text{ X } S \rightarrow [0,1]$$

If you knew which state of the world is actual, you could use A to calculate the accuracy of your credence function. But of course you don't know which state is actual, and so you're not in a position to know how accurate you are. Nevertheless, you are in a position to calculate your *expected accuracy*. The expected accuracy of a credence function $c \in C_s$, relative to a probability function $p \in C_s$, is defined as follows:

$$\mathbf{EA}_p(c) = \sum_{s \in S} p(s)A(c,s)$$

The expected accuracy of a credence function c relative to p is the average of the accuracy scores c would get in the various states that might obtain, weighted by how likely those states are to obtain according to p.

A *strictly proper* scoring rule is a scoring rule with the feature that every probabilistic credence function maximizes expected accuracy relative to itself. More formally, where P_s is the set of probabilistic credence functions defined over S, A is a strictly proper scoring rule just in case:

$$\text{For every } p \in P_s \text{ and every } c \in C_s \text{ such that } p \neq c, \mathbf{EA}_p(p) > \mathbf{EA}_p(c)$$

When subjects use strictly proper scoring rules to measure the accuracy of their credence function, they are *strongly immodest*. They expect their own credences to be more accurate than any others. (A subject is *mildly immodest* if she expects her credences to be at least as accurate as any others.)

Strong Immodesty has the potential to solve both of our problems. If you are immodest, you can never be uncertain about which attitude best satisfies your aim of believing accurately—you will always expect your own credence to be more accurate than every other credence. Strong Immodesty can also help us explain why it's impossible to raise or lower your credence at will. If we assume you have a precise credence in the proposition that God exists, and that you're strongly immodest, then changing your credence on the basis of a coin flip would mean adopting a credence that is, by your own lights, less accurate than your own.

But it's not just that Strong Immodesty *can* solve our two problems. It's that we have no hope of solving them without it. Recall the spelling analogy. If you aren't sure whether 'cemetery' is spelled 'C-E-M-E-T-E-R-Y' or 'C-E-M-E-T-A-R-Y,' you can choose to spell it either way. It's true that your aim is to spell 'cemetery' correctly, but from your perspective, there are two equally good ways to satisfy that aim. That is why you have some discretion over how you spell the word. Without Strong Immodesty, belief is just like spelling. If your credence in P is not strongly immodest, then there is some other credence in P that you expect to be no less accurate than your own. Both satisfy your aim of believing accurately, and so we can't explain why you can't choose which to adopt on the basis of a coin flip. The success of the aim-of-belief argument depends on whether we can show that it is part of having a credence—not just part of having a *rational* credence—that you expect it be more accurate than any other. We need to show that Strong Immodesty is a constitutive requirement, not just a rational requirement.

But Strict Propriety does not say that just *any* credence function expects itself to be more accurate than any other credence function. It says that any probabilistically coherent credence function—any credence function that obeys to axioms of the probability calculus—expects itself to be more accurate than any other credence function. Probabilistically incoherent credences aren't self-recommending.

What should we say to this objection? One option is to argue that incoherent agents are *fragmented*. Consider Mary, who appears to be quite confident in p and quite confident in $\neg p$. Rather than say that Mary really does assign, say, .7 to p and .7 to $\neg p$, we say that Mary doesn't have a single incoherent system of beliefs—either binary or graded—but many different, compartmentalized systems of belief, which guide different aspects of her behavior in different contexts. Importantly, each of Mary's systems of belief—each of her fragments—is probabilistically coherent. One fragment assigns .7 to p, and .3 to $\neg p$; another assigns .7 to $\neg p$ and .3 to p.

If Mary's doxastic state is fragmented, we can't say that she uses a strictly proper scoring rule to measure the accuracy of her credence function, since she doesn't have a single credence function. What we can say, however, is that each of

Mary's *fragments* uses a strictly proper scoring rule to measure the accuracy of *its* credence function. And maybe that is enough. If each of Mary's fragments expects its own credences to be most accurate, and each aims at believing accurately, then changing its credences on the basis of a coin flip would mean adopting a credence that is, by its own lights, less accurate than its own. And that, the argument continues, is impossible. Just as one cannot intentionally form a belief unless one is doing so with the aim of forming a true belief, one cannot intentionally increase or decrease one's confidence unless one is doing so with the aim of increasing one's accuracy.

There's a lot more to say about this. But I propose that we move on. Even if we can use fragmentation to explain why probabilistically incoherent subjects cannot believe at will, the aim-of-belief argument faces another serious objection. In the next section, I will suggest that this problem—the *problem of imprecision*, as I will call it—may well be insurmountable.

3.4 The Problem of Imprecision

So far we've assumed that credal states are best modeled by a single probability function, a function that assigns to each proposition a unique, precise credence. For many Epistemic Bayesians, this is a crude idealization, demanding far more precision than is usually warranted by our messy evidence. These *Imprecise Epistemic Bayesians* propose to model agents not with one probability function, but many—a *set* of credence functions; a kind of credal committee.[10] Each member of your credal committee is a perfect Bayesian agent, assigning a single, precise credence to every proposition. But not you. Your attitude may be 'spread out' over a range of values.

Take the proposition that you'll get sick in the next 6 months. You probably don't have a precise credence in this claim, the Imprecise Bayesian says. Instead, you're imprecise—say, 60%-to-90% confident that you'll catch something within six months. We model you with a set of probability functions: for every number

[10]For classic discussions of imprecise credences, see Kyburg (1983) and Levi (1974). See Joyce (2005), Joyce (2011), Moss (2014), Schoenfield (2012), and Sturgeon (2008) for contemporary defenses.

in the interval [.6, .9], some probability function—some member of your credal committee—assigns that credence to the proposition that you will get sick in the next six months.[11]

Sets of probability functions reflect degrees of evidential *specificity*. As Joyce (2005) points out, there are different ways for the evidence to be unspecific. It can be *incomplete* in the sense that it doesn't discriminate one hypothesis from alternative hypotheses, and it can be *ambiguous*, in the sense that it is subject to different interpretations. (Of course both incompleteness and ambiguity are matters of degree.) When your credence in a certain proposition P is imprecise, you have no settled opinion about it. The members of your credal committee represent various, more settled views on P that you might take, corresponding to different interpretations of the evidence (in the case of evidential ambiguity) or different ways of 'filling in the gaps' (in the case of incompleteness).

Imprecise credences present a serious problem for the aim-of-belief argument. For it is hard to see how subjects with imprecise credences could be immodest. There seems to be no plausible way of scoring imprecise credences that makes them self-recommending. The formal arguments for this claim can be rather nuanced, however, and I haven't the space to go into the details. I will be content to offer a more abstract and intuitive—but, I must admit, much less precise—explanation.[12]

There are two ways to score imprecise credences. Either we score them *precisely*—with a real number between 0 and 1, in just the way we score precise credences. Or we score them *imprecisely*—with a set of numbers.

Beginning with the precise strategy, I can think of only two principled ways to score imprecise credences precisely. We could identify the accuracy of [.4, .6] at a world with the accuracy of the *midpoint* of the interval at that world—in this case, the accuracy of .5. Or we could identify the accuracy of [.4, .6] with the *average* of the accuracy scores of each of the members of the interval. We consider each credence in the interval in turn, calculate its accuracy at a given world, then take

[11]Plausibly, the edges of these ranges are vague. I ignore this complication, since nothing I say turns on it.

[12]For more detailed arguments see Schoenfield (2017) and Seidenfeld, Schervish, and Kadane (2012).

the average of those values.

First consider the midpoint rule. If we identify the accuracy of [.4, .6] with the accuracy of .5, then [.4, .6] will be no more accurate than .5 *at every world.* But it is hard to see how [.4, .6] could recommend itself over. .5 if, at every world, .5 is just as accurate as [.4, .6].

The second precise proposal is no better off. If we identify the accuracy of the imprecise credence with the average of the accuracy scores of the precise members of the interval, we can prove that, for any imprecise state S there is some precise state S' that is more accurate than S in every world—in other words, every imprecise state will be accuracy dominated by some precise state.[13] But again, if S' is more accurate than S is every world, S will not recommend itself over S'.

Finally, turn to the imprecise method of scoring imprecise credences. On this proposal, we measure the score of an imprecise state with a set of numbers—most obviously, with the set of accuracy scores of the precise credences in the interval. Now, if the accuracy of an imprecise credence, at a given world, is a set of numbers, and the accuracy of a precise credence at that world, is a single number, then plausibly their accuracy scores are *incomparable*. This means that, at every world,[.4, .6] is no more accurate than .5, .5 is no more accurate than [.4, .6], and they aren't equally accurate, either. But once again, it's hard to see how [.4, .6] could recommend itself over .5 if, at every world, [.4, .6] is no more accurate than .5.

We considered three ways to score imprecise credences. None of them ensures imprecise credences are self-recommending. For any imprecise credal state, we can always find a precise state whose accuracy, at every world, is no worse than the accuracy of the imprecise state. But that's just to say that you know, of some particular precise state, that it is no less accurate than your current imprecise state. And if that's right, we can't rule out changing one's credence at will. If there are two distinct credal states, such that you know that neither is less accurate than the other, then, it seems, adopting either will satisfy your aim of maximizing accuracy. No reason, then, not to choose by flipping a coin.

[13] This follows from a much more general result from Mayo-Wilson and Wheeler (2016).

3.5 Conclusion

I've considered two arguments for the seeming impossibility of believing at will, both inspired by Bernard Williams's suggestive remarks on the matter—the retrospective argument, and the aim-of-belief argument, respectively. I have shown that neither argument succeeds in showing that it is impossible to believe at will. The retrospective argument allows that you can decide to believe in acknowledged permissive cases, those in which you believe that your evidence permits believing and permits not believing. The aim-of-belief argument allows that we can change our doxastic states at will when they are imprecise.

References

1. Alchourrón, C.E., P. Gärdenfors, and D. Mackinson (1985). 'On the Logic of Theory Change: Partial Meet Contraction and Revision Functions.' *Journal of Symbolic Logic*, 50: 510-530.

2. Ballantyne, Nathan and Coffman, E.J. 2012, 'Conciliationism and Uniqueness'. *Australasian Journal of Philosophy*, 90, pp. 657–70.

3. Buchak, Lara (2014). 'Belief, Credence, and Norms.' *Philosophical Studies*, 169 (2): 285-311.

4. Christensen, David 2010, 'Rational Reflection'. *Philosophical Perspectives*, 24, pp. 121–40.

5. Christensen, David and Lackey, Jennifer (eds.) 2013, *The Epistemology of Disagreement: New Essays*, Oxford: Oxford University Press.

6. Clarke, Roger (2013). 'Belief is Credence One (in Context).' *Philosophers' Imprint*, 13 (11): 1-18.

7. Coates, Allen 2012, 'Rational Epistemic Akrasia'. *American Philosophical Quarterly*, 49, pp. 113–24.

8. Cohen, Stewart 2013, 'A Defense of the Almost Equal Weight View'. In Christensen and Lackey, 2013, pp. 98–119.

9. Dorst, Kevin (forthcoming). 'Lockeans Maximize Expected Accuracy.' *Mind*.

10. Dougherty, Trent (ed.) 2011, *Evidentialism and its Discontents*. Oxford: Oxford University Press.

11. Douven, Igor. 2009, 'Uniqueness Revisited'. *American Philosophical Quarterly*, 46, pp. 347–61.

12. Elga, Adam. 2013, 'The Puzzle of the Unmarked Clock and the New Rational Reflection Principle'. *Philosophical Studies*. 164, pp. 98-120.

13. Frankish, Keith (2007). 'Deciding to Believe Again.' *Mind* 116 (463): 523-547.

14. Ginsberg, Matthew (1986). 'Counterfactuals.' *Artificial Intelligence*, 30.

15. Goodman, Jeremy and Salow, Bernhard (forthcoming). 'Taking a Chance on KK.' *Philosophical Studies.*

16. Greaves, Hilary and Wallace, David (2006). 'Justifying Conditionalisation: conditionalisation maximizes expected epistemic utility.' *Mind*, 115 (459): 607-632.

15. Greco, Daniel 2014, 'A Puzzle about Epistemic Akrasia'. *Philosophical Studies*. 167, pp. 201–19.

16. Greco, Daniel and Hedden, Brian (2016). 'Uniqueness and Metaepistemology.' *Journal of Philosophy* 113 (8): 365-395.

17. Hall, Ned (1999). 'How to Set a Surprise Exam.' *Mind*, 108 (432): 647-703.

18. Hawthorne, John, Rothschild, Daniel, and Spectre, Levi (2016). 'Belief is Weak.' *Philosophical Studies*, 173 (5): 1393-1404.

19. Horowitz, Sophie (2014). 'Immoderately Rational.' *Philosophical Studies* 167 (1): 41-56.

20. Horowitz, Sophie (2014). 'Epistemic Akrasia'. *Noûs*, 48, pp. 718-44.

21. Joyce, James (2005). 'How Probabilities Reflect Evidence.' *Philosophical Perspectives*, 19 (1): 153-178.

22. Joyce, James (2011). 'A Defense of Imprecise Credence in Inference and Decision-Making'. *Philosophical Perspectives*, 24 (1): 281-323.

23. Keynes, John Maynard (1921). *A Treatise on Probability*. Dover Publications.

24. Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitrechnung*. Ergebnisse Der Mathematik; translated as *Foundations of Probability*, New York: Chelsea Publishing Company, 1950.

25. Kopec, Matthew 2015, 'A Counterexample to the Uniqueness Thesis'. *Philosophia*, 43, pp. 403-09.

25. Kopec, Matthew and Titelbaum, Michael G. 2015, 'The Uniqueness Thesis'. forthcoming in *The Philosophical Compass*.

26. Kyburg, Henry (1983). 'Rational Belief'. *Brain and Behavioral Sciences.* 6 (2): 231-273.

27. Lasonen-Aarnio, Maria 2014, 'Higher-order Evidence and the Limits of Defeat'. *Philosophy and Phenomenological Research*, 88, pp. 314–45.

28. Lasonen-Aarnio, Maria 2015, 'New Rational Reflection and Internalism about Rationality'. In Hawthorne and Gendler 2015, pp. 145-71.

29. Lawlor, Krista (2014). 'Exploring the Stability of Belief: Resiliency and Temptation'. *Inquiry*, 57 (1): 1-27.

30. Lehrer, Keith (1971). 'How Reasons Give Us Knowledge or the Case of the Gypsy Lawyer.' *Journal of Philosophy* 68 (10): 311-313.

31. Leitgeb, Hannes (2013). 'Reducing Belief Simpliciter to Degrees of Belief.' *Annals of Pure and Applied Logic*, 164: 1338-89.

32. Leitgeb, Hannes (2014). 'The Stability Theory of Belief'. *Philosophical Review*, 123 (2): 131-171.

33. Leitgeb, Hannes (2017). 'The Stability of Belief: How Rational Belief Coheres with Probability'. Oxford University Press.

34. Levi, Isaac (1974). 'On Indeterminate Probabilities'. *Journal of Philosophy*, 71 (13): 391-418.

35. Loeb, Louis (2001). 'Integrating Hume's Accounts of Belief and Justification.' *Philosophy and Phenomenological Research*, 63(2): 279-303.

36. Loeb, Louis (2002). *Stability and Justification in Hume's Treatise*. Oxford University Press.

37. Kevin McCain & Poston, Ted. 'Why Explanatoriness is Evidentially Relevant.' *Thought: A Journal of Philosophy*, 3 (2): 145-153.

38. Meacham, Christopher J.G. 2014, 'Impermissive Bayesianism'. *Erkenntnis*, 79, pp. 1185–1217.

39. Moss, Sarah (2014). 'Credal Dilemmas'. *Noûs* 48 (3): 665-683.

40. Rosen, Gideon 2001, 'Nominalism, Naturalism, and Epistemic Relativism'.

Philosophical Perspectives, 15, pp. 69–91.

41. Salow, Bernhard (ms). 'Fallibility and Dogmatism.'

42. Schoenfield, Miriam (2012). 'Decision Making in the Face of Parity.' *Philosophical Perspectives* 28 (1): 263-277.

43. Schoenfield, Miriam (2013). 'Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief.' *Noûs* 47 (1): 193-218.

44. Schoenfield, Miriam (2017). 'The Accuracy and Rationality of Imprecise Credences.' *Noûs* 51 (4): 667-685.

45. Schoenfield, Miriam (forthcomiming). 'Permissivism and the Value of Rationality: A Challenge to the Uniqueness Thesis.' *Philosophy and Phenomenological Research.*

46. Setiya, Kieran. 'Believing at Will.' *Midwest Studies in Philosophy* 32 (1): 36-52.

47. Setiya, Kieran (2013). 'Epistemic Agency: Some Doubts.' *Philosophical Issues* 23 (1): 179-198.

48. Shah, Nishi and Velleman David (2005). 'Doxastic Deliberation.' *Philosophical Review* 114 (4): 497 - 534.

49. Skyrms, Brian (1977). 'Resiliency, Propensities, and Causal Necessity'. *Journal of Philosophy*. 74 (11): 704-713.

50. Sliwa, Paulina and Horowitz, Sophie (2015). 'Respecting all the Evidence'. *Philosophical Studies*, 172, 2835–2858.

51. Smithies, Declan (2012). 'Moore's Paradox and the Accessibility of Justification'. *Philosophy and Phenomenological Research*, 85, pp. 273–300.

52. Sturgeon, Scott (2008). 'Reason and the Grain of Belief'. *Noûs*, 42 (1): 139-165.

53. Szabó Gendler, Tamar and Hawthorne, John (eds.) 2015. *Oxford Studies in Epistemology*, volume 5. Oxford: Oxford University Press.

54. Teller, Paul (1976). 'Conditionalization, Observation, and Change of Preference' in Harper, W. and Hooker, C.A. (eds.) *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science.* D. Riedel.

55. Titelbaum, Michael (2010). 'Not Enough There There: Evidence, Reasons, and Language Independence'. *Philosophical Perspectives*, 24, pp. 477-528.

56. Titelbaum, Michael (2015a). 'Continuing On'. *Canadian Journal of Philosophy*, 45, pp. 670–691.

57. Titelbaum, Michael (2015b). 'Rationality's Fixed Point (or: In Defense of Right Reason)'. In Gendler and Hawthorne 2015, pp. 253–294.

58. van Fraassen, Bas C. (1989). *Laws and Symmetry*. Oxford University Press.

59. van Fraassen, Bas C. (1999). 'A New Argument for Conditionalization.' *Topoi*, 18: 93-96.

60. Weatherson, Brian (2005). 'Can We Do Without Pragmatic Encroachment?' *Philosophical Perspectives*, 19 (1): 417-443.

61. White, Roger (2005). 'Epistemic Permissiveness.' *Philosophical Perspectives* 19: 445-459.

62. Williams, Bernard (1981). 'Deciding to Believe.' in *Problems of the Self*, 136-151. Cambridge, England: Cambridge University Press.

63. Williams, Peter F. (1980). 'Bayesian Conditionalisation and the Principle of Minimum Information', 31(2): 131-144.

64. Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

65. Williamson, Timothy (2001). 'Improbable Knowing'. In Trent Dougherty, 2011, pp. 149–64.