Evidence as a Guide to Truth

by

Sophie Horowitz

Submitted to the Department of Linguistics and Philosophy on May 21, 2014 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Philosophy.

Abstract

Any non-skeptical epistemological theory should hold that, if we want to know what the world is like, our best bet is to believe what our evidence supports. Evidence is a guide to truth. Though it might sound like a platitude, this thesis is far from trivial. My dissertation defends and develops a more complete picture of what our epistemological theory should look like, given this anti-skeptical thesis.

If evidence is a guide to truth, we should expect that our beliefs about what our evidence supports should line up with our beliefs about the world. But this idea has come under fire in recent literature. In the first chapter, “Epistemic Akrasia”, I argue that it is irrational to believe both P and my evidence does not support P. I show that epistemic akrasia licenses bad reasoning and irrational action.

The second chapter “Immoderately Rational”, brings out a surprisingly strong implication of the thesis. I argue that we can defend the thought that evidence is a guide to truth only if we have an extremely impermissive view, on which there is just one ideally rational response to any body of evidence. Even moderately permissive views cannot give this answer.

In the third chapter, “Immodesty and Educated Guesses”, I step back to consider how degrees of belief, or credences, can “get things right”, given that they cannot be true or false. I defend a novel alternative: credences get things right by licensing true “educated guesses”. The idea is that an agent’s credences license all-or-nothing verdicts, which can be elicited in forced choice scenarios; credences get things right to the extent that these all-or-nothing verdicts tend to be true. This account vindicates a popular and plausible thought that rational credence is stable: if you know your credences are rational, you should not want to change them without receiving new evidence. I also suggest that it can be used to argue for probabilistic coherence as a rational requirement.
Acknowledgments

For helpful comments, discussion, questions, and suggestions, thanks to many, many people: Roger White, Steve Yablo, Alex Byrne, David Christensen, Brian Hedden, Jack Marley-Payne, Bernhard Salow, Miriam Schoenfield, Katia Vavova, Alejandro Pérez Carballo, Paulina Sliwa, Ryan Doody, Daniel Greco, Brian Weatherson, Adam Elga, Josh Schechter, Richard Holton, Kenny Walden, Tom Dougherty, Chris Meacham, Louis deRosset, Tyler Doggett, Sinan Dogramaci, Caspar Hare, Alex Silk, Justin Khoo, Richard Pettigrew, Agustin Rayo, Mike Titelbaum, Fritz Warfield, Declan Smithies, Scott Sturgeon, Rachael Briggs, Jennifer Carr, Matthias Jenny, Jonathan Vogel, Robert Stalnaker, Nilanjan Das, Brendan Dill, David Gray, Sally Haslanger, Ned Markosian, Jack Spencer, Kevin Richardson, Brad Skow, Dennis Whitcomb, Ryan Wasserman, Damien Rochford, Daniel Hagen, Julia Markovits, Ian Wells, Arden Ali, Rose Lenehan, Rebecca Millsop, Matt Mandelkern, Melissa Schumacher, and an anonymous referee for *Noûs*.

This work has also been presented, in some form or another, at several different places. Thanks to audiences at the 2011 Columbia/NYU Graduate Conference, the 2013 UT Austin Graduate Conference, the 2013 Notre Dame/Northwestern Graduate Epistemology Conference, the 2013 Bellingham Summer Philosophy Conference, and to my commentators: Katrina Przyjemski, Sam Fox Krauss, Katie Finley, Kenny Easwaran, and Miriam Schoenfield. Thanks also to audiences and readers at many amazing presentation venues at MIT: WIP, ERG, MATTI, and DOODY.

Special thanks to David Christensen, Miriam Schoenfield, Paulina Sliwa, Katia Vavova, Bernhard Salow, and Brian Hedden for extended discussion of many, many drafts of some chapters, which was far beyond the call of duty. And extra special thanks to my committee — Alex Byrne, Steve Yablo, and Roger White — for extended discussion of many, many drafts of everything. Roger White, my advisor, was extremely helpful and generous every step of the way.

Thanks to Paulina Sliwa and Katia Vavova for cooking me things, and to Paulina, Katia, Tom Dougherty, and Kenny Walden and being excellent official and unofficial big buddies. Thanks to 1369 and Mariposa Bakery for letting me sit for a long time without buying very much. Thanks to our forthcoming baby for making this past year more interesting and more sober than it otherwise would have been. Finally, thanks to Ruth Horowitz, David Christensen, Sam Horowitz, and most especially Henry Swift, for their love and support.
Chapter 1
Epistemic Akrasia

Many views rely on the idea that it can never be rational to have high confidence in something like, “P, but my evidence doesn’t support P.” Call this idea the “Non-Akrasia Constraint.” Just as an akratic agent acts in a way she believes she ought not act, an epistemically akratic agent believes something that she believes is unsupported by her evidence. The Non-Akrasia Constraint says that ideally rational agents will never be epistemically akratic.¹

The importance of the Non-Akrasia Constraint is brought out by recent literature on “higher-order evidence” – evidence about what evidence one has, or what one’s evidence supports. Much of this debate has focused on the question of what to believe when one becomes rationally highly confident that P on the basis of some evidence, E, and then receives further (misleading) evidence to the effect that E does not support P. Although there is disagreement as to whether this new evidence should cause one to reduce confidence in P, the major positions that have been defended thus far in the higher-order evidence debate agree that ideally rational agents should respect the Non-Akrasia Constraint.

But in a number of recent papers, the Non-Akrasia Constraint has been called into question. Some have argued that in cases of misleading higher-order evidence, epistemic akrasia can in fact be rationally required. If you have excellent evidence for P, and excellent evidence that your evidence doesn’t support P, the thought goes, perhaps the

¹Some authors take the irrationality of epistemic akrasia as a premise. See, e.g., Smithies [2012] and Titelbaum [ms]. Feldman [2005] defends a broadly anti-akratic view, and writes that “[o]ne wonders what circumstances could make [epistemic akrasia] reasonable” (p. 108-9). Christensen [2007 b] argues that it’s not rational to be certain of logical theorems while doubting that you are perfectly reliable at proving these theorems (p. 18-19). Many arguments in the literature on disagreement and higher-order evidence rely on the idea that if it is rational for you to believe that a certain belief of yours is unjustified, you should reduce confidence in that belief. See, e.g., Christensen [2007 a], [2010], and [2011]; and Elga [2007] and [ms]. Kolodny [2005] assumes that epistemic akrasia is a type of internal conflict, and suggests that one’s evidence will never support being in such a state. Adler [2002] argues that it is psychologically impossible to believe P, be fully aware that one believes P, and also believe that one’s belief in P is unsupported by one’s evidence. Greco [forthcoming] argues that we should understand akratic belief states as “fragmented”, and that this justifies our taking them to be irrational.

For variety’s sake, I’ll sometimes describe akratic states as beliefs or high confidence in conjunctions, and sometimes as combinations of attitudes. Strictly speaking, the Non-Akrasia Constraint should be understood simply as a prohibition on having high confidence in both “P” and “my evidence doesn’t support P”.

6
rational response is to have high confidence in both of these things. The tenability of these pro-akrasia (or “Level-Splitting”) views will depend, in part, on whether there is a strong case for the Non-Akrasia Constraint. So what can be said in favor of the constraint?

The goal of this paper is to defend the Non-Akrasia Constraint... for the most part. In the first part of the paper I defend the constraint indirectly, by showing that Level-Splitting views face serious challenges: they have counterintuitive consequences regarding the rational response to higher-order evidence, and it is unclear how these views could deliver plausible results in nearby cases. The upshot of this discussion, and of the bulk of the paper, is that we should accept the Non-Akrasia Constraint.

In the last section, I consider a different kind of challenge to Non-Akrasia, based on Williamson’s case of the unmarked clock. Looking at this less-paradigmatic type of case gives us reason to think that Non-Akrasia may not be a universal rational requirement. But, surprisingly, I suggest that allowing rational epistemic akrasia under certain special conditions does not undermine anti-akratic verdicts (or the arguments against Level-Splitting) in paradigm cases.

1. A Paradigm Case

Consider the following story:

Sleepy Detective: Sam is a police detective, working to identify a jewel thief. He knows he has good evidence – out of the many suspects, it will strongly support one of them. Late one night, after hours of cracking codes and scrutinizing photographs and letters, he finally comes to the conclusion that the thief was Lucy. Sam is quite confident that his evidence points to Lucy’s guilt, and he is quite confident that Lucy committed the crime. In fact, he has accommodated his evidence correctly, and his beliefs are justified. He calls his partner, Alex. “I’ve gone through all the evidence,” Sam says, “and it all points to one person! I’ve found the thief!” But Alex is unimpressed. She replies: “I can tell you’ve been up all night working on this. Nine times out of the last ten, your late-night reasoning has been quite sloppy. You’re always very confident that you’ve found the culprit, but you’re almost always wrong about what the evidence supports. So your evidence probably doesn’t support Lucy in this case.” Though Sam hadn’t attended to his track record before, he rationally trusts Alex and believes that she is right – that he is usually wrong about what the evidence supports on occasions similar to this one.
What should Sam believe about who the jewel thief is? And what should he believe about what his evidence supports?

In Sleepy Detective, Sam’s epistemic position can be described in the following way: first, he receives some “first-order” evidence, \( E \), which includes the letters and photographs that he was looking through as he worked late at night. \( E \) supports a proposition, \( P \): that Lucy is the thief. So after considering \( E \), it seems that Sam should believe \( P \). Then, when Sam talks to Alex, he receives some higher-order evidence, \( HOE \): her testimony about his track record data, which suggests that in situations like this one, he almost always misinterprets his evidence. \( HOE \) therefore suggests that the answer Sam came up with is unlikely to be supported by his evidence; since Sam came up with \( P \), \( HOE \) suggests that \( P \) is unlikely to be supported by Sam’s evidence. (Note that in order for this to be true, we don’t have to suppose that Sam is in fact unreliable; all we need is that \( HOE \) supports the proposition that he is.) So \( HOE \) supports this proposition: that \( E \) does not support \( P \).

What does Sam’s total evidence support?

Two answers to this question have been most commonly defended in the literature; I will mention them only briefly, and then largely set them aside for the remainder of the paper. According to “Steadfast” views, Sam should maintain his high confidence in \( P \), and also maintain his high confidence that \( E \) supports \( P \). The advantage of Steadfast views is that they acknowledge the evidential force of \( E \); the disadvantage is that they seem to ignore \( HOE \). According to “Conciliatory” views, Sam should become much less confident in \( P \), and also become less confident that \( E \) supports \( P \). The advantage of Conciliatory views is that they acknowledge the evidential force of \( HOE \); the disadvantage is that they seem to ignore \( E \). Although Steadfast and Conciliatory views disagree about what Sam’s total evidence supports, they share a common assumption: that Sam’s confidence in the first-order proposition, \( P \), should line up with

---

2 See Kelly [2005] and Titelbaum [ms] for defenses of this type of view, applied specifically to the case of peer disagreement.

3 See Feldman [2005] for a view like this; see also the recent literature on disagreement and higher-order evidence by, e.g., Christensen and Elga. Kelly [2007]’s “Total Evidence” view could be read as a compromise between Steadfast and Conciliatory views that still respects the Non-Akrasia Constraint. So, even among broadly Conciliatory views, just how much Sam should decrease confidence is up for debate. Sleepy Detective is meant to parallel cases discussed by all of these authors regarding higher-order evidence.
his confidence in the proposition that E supports P. Thus, both Steadfast and Conciliatory views respect the Non-Akrasia Constraint in cases like Sleepy Detective.

The view that I will discuss here, which has become more prominent in recent literature, offers a third answer. Call this third view “Level-Splitting”. Level-Splitting says that Sam should maintain his high confidence in P, but he should become much less confident that E supports P. In short, he should become highly confident in “Lucy is the thief, but my evidence doesn’t support that.” Level-Splitting might seem like a happy compromise between Steadfast and Conciliatory views. In accommodating both first-order and higher-order evidence, it has the apparent advantages of both views, and the apparent disadvantages of neither. But in one important respect, Level-Splitting is not a compromise between Steadfast and Conciliatory views. While Steadfast and Conciliatory views both respect the Non-Akrasia Constraint, Level-Splitting denies it.

In the next section, I will turn to discussing some Level-Splitting views that have been defended in recent literature.

2. Level-Splitting views

The view I’m calling “Level-Splitting” will be picked out by its verdict in cases like Sleepy Detective: that it can be rational, in that type of case, to have high confidence in both “P” and “my evidence doesn’t support P.” Of course, there are many possible ways to motivate this verdict, and the four authors I consider below arrive at the view in different ways. Looking at these views will shed light on what kinds of epistemological positions are compatible with the Non-Akrasia Constraint.

2.1 Williamson

In his “Improbable Knowing” and “Very Improbable Knowing”, Timothy Williamson argues that there can be cases where one knows that P, but it is also highly improbable on one’s evidence that one knows that P. In certain cases, Williamson’s view is committed to a denial of the Non-Akrasia Constraint: it requires a rational agent to have high confidence in “P, but my evidence doesn’t support P.”

One example of this type is what Williamson calls the “Long Deduction”. Suppose a rational agent comes to know a series of claims and competently deduces their
conjunction, C. On Williamson’s view, she can come to know C by these means. But suppose further that the rational agent also knows that, since most people’s memories and logical abilities are limited, people in her situation often make inferential errors while completing long deductions. It can then be highly probable on her evidence that she has made such an error, and thus that she does not know the conjunction. Nevertheless, since she has in fact competently deduced C, Williamson holds that she does know C. In a case like this, it will be certain on her evidence that C is true – for Williamson, knowledge of C requires assigning C evidential probability 1 – but also highly probable on her evidence that she does not know C. So she should be highly confident in C, and highly confident that she does not know that C.

High confidence in “C” and “I don’t know that C” is strange, but it’s not yet a violation of the Non-Akrasia Constraint. (“I don’t know that C” is compatible with “C is highly probable on my evidence”, and high confidence in “C, and C is highly probable on my evidence” does not have quite as much of a Moore-paradoxical flavor. I’ll say more about less-extreme kind of mismatch in section 2.5.) On a plausible filling out of Williamson’s story, however, an ideally rational agent can also be highly confident of both “C” and “my evidence doesn’t support C”.

We have seen that our rational agent can be rationally highly confident that she does not know C because of doubts about inferential mistakes; so she must be rationally highly confident that she has made some such mistake. What kind of mistake should she think she has made? Usually, when we make mistakes in completing long deductions, we reach false conclusions. But our rational agent can’t think that this is what’s happened; after all, she knows that the conclusion she reached is C, and she is certain that C is true. So she must think she has made some other kind of inferential error: one that would result in a true belief, but not knowledge.

Whether a rational agent will become highly confident of both “C” and “my evidence doesn’t support C”, in these circumstances, will depend on what kind of inferential error she thinks she has made. Some kinds of error are compatible with C’s

4 See Williamson [notes a], p. 6, and [notes b], p. 36-42, for discussion of this kind of case. In his [notes b], fn 19, Williamson explicitly discusses cases where one’s lack of higher-order knowledge comes about because of doubts that one’s deduction is free from “inferential error”.

5 Thanks to an anonymous Nous referee for helpful suggestions regarding this point.
being supported by her evidence. For example, if it's a mathematical deduction, she might think she made several calculation errors that canceled one another out, luckily resulting in a true belief. Or she might worry that, even though her evidence (whatever it is) supports C, there's a good chance that her belief isn't properly "based on" that evidence. In these cases, the agent should believe C is true, and supported by her evidence – but nevertheless she might be rationally highly confident that her belief in C does not amount to knowledge.

Errors of a second type would make it likely that C is not supported by her evidence. For example, she might think she has misidentified her evidence, and that the evidence she could have mistaken it for does not support C. Or she might think she has made a mistake in identifying what her evidence supports: although her evidence doesn't support C, C is true, and she (luckily, but irrationally) believes it. If an agent has reason to believe that she has made an error of this second type, it is plausible that she should become highly confident that her evidence doesn't support C.

Williamson agrees that there can be cases in which an agent is unsure of what her evidence is, or of what her evidence supports. And it is possible that sometimes one can be not only unsure, but significantly mistaken: as Williamson writes in his "Very Improbable Knowing", "one's evidence can be radically misleading about one's own present epistemic position." In other cases, Williamson allows that the evidential probability of P can be quite high, while it is also highly probable that the evidential probability of P is quite low. So it's plausible that, on his view, there can be cases where an agent is rationally highly confident that C (as a result of competent deduction) but also rationally highly confident that her evidence doesn't support C (because she is rationally highly confident that she has made an error of the second type above).

One such case might be Sleepy Detective: while it's highly likely on Sam's first-order evidence that Lucy is the thief, it's also highly likely on Sam's higher-order evidence that that he has made some kind of error in evaluating his evidence. If Sam has

---

6 Williamson [notes b], p. 2.
7 See Williamson (2000), p. 229. Williamson describes a case in which P is 80% probable on one's evidence, but it is 80% probable that P is only 20% probable. This example is a useful illustration of one of the many ways in which Williamson's view allows one's rational confidence in "P" and "my evidence supports P" to come apart quite widely when one has reason to be uncertain about one's evidence. This particular case deserves further attention, and I will discuss it in detail in Section 6. For present purposes, however, I will focus on cases like Sleepy Detective.
reason to believe that his error is of the more problematic type – one that would lead him to believe something unsupported by his evidence – it seems that on Williamson’s view, Sam should be highly confident of both “Lucy is the thief” and “my evidence doesn’t support that.”

2.2 Lasonen-Aarnio

Maria Lasonen-Aarnio, in her “Higher-Order Evidence and the Limits of Defeat”, also argues that one can have rational high confidence in both “P” and “my evidence doesn’t support P” in cases of misleading higher-order evidence. Lasonen-Aarnio’s reasons for thinking so arise from a puzzle that emerges when one tries to articulate precisely how higher-order evidence could defeat first-order attitudes. Prima facie, she points out, it appears that higher-order defeat requires rationality to give “conflicting recommendations”. In Sleepy Detective, for example, one correct epistemic rule apparently tells Sam to be highly confident in P (since arguably, that’s what it is for Sam’s first-order evidence to strongly support P). And if higher-order defeat is a genuine phenomenon, there is another correct epistemic rule telling Sam to reduce confidence in P. So Sam is subject to contradictory norms. How could this be?

Lasonen-Aarnio surveys several possible responses to the puzzle, and argues that none is satisfactory. Ultimately, Lasonen-Aarnio suggests that there can be no single coherent, non-paradoxical notion of rationality that accommodates higher-order defeat.8

What conclusion should we draw? Lasonen-Aarnio argues that we should give up on the idea of higher-order defeat: sometimes a rational agent can accommodate higher-order evidence in her attitudes about what her evidence supports without revising her confidence in her first-order attitudes. Lasonen-Aarnio writes that “in so far as there is such a thing as a correct inductive policy or epistemic system, it can be rational to follow the recommendations of that policy or system even if one possesses evidence that in doing so one has committed a rational error;” and “[t]hat one should believe that one

8 This is not to say that there is nothing wrong, epistemically speaking, with someone who maintains confidence in P after receiving evidence that her belief in P is unjustified or unsupported by her evidence. Lasonen-Aarnio suggests that there is a separate epistemic norm, “reasonableness,” that makes this kind of belief state criticisable too. See also her [2010] for more on this idea.
shouldn’t Φ doesn’t entail that one shouldn’t Φ.’’ So in rejecting higher-order defeat in the way she does, Lasonen-Aarnio also rejects the Non-Akrasia Constraint.

2.3 Wedgwood
A third example of a recent view along the lines of Level-Splitting comes from Ralph Wedgwood’s “Justified Inference”. On Wedgwood’s view, an inference is justified if and only if, in drawing that inference, one actually manifests a “necessarily rational disposition”. He writes,

> [I]t is the real nature of this internal process—the fact that it is a process of competent inference—and not the higher-order beliefs that the thinker has, or even the beliefs that the thinker is justified in having, about the nature of that process, that is crucial to the rationality of the mental event that results from that process.  

Beliefs formed through justified inferences will also be justified. And if higher-order evidence doesn’t change the “real nature” of one’s reasoning process, it also does not rationally mandate that one revise one’s confidence in the first-order beliefs formed using that process.

In discussing higher-order evidence, Wedgwood writes that on his view, “you might be in the weird situation of rationally drawing an inference even though it is not rational for you to believe the inference to be rational.” This is true even when you have reason to believe that the inference was not rational: Wedgwood writes, “a perfectly rational thinker would continue to draw the inference even if she had (misleading) evidence that she was reasoning incompetently, and even if she entertained serious doubts about whether or not she really was perfectly rational.”

So in Sleepy Detective, if Sam’s inference to the conclusion that Lucy is the thief was in fact an instance of the right kind of mental process, nothing Alex could tell Sam about his own rationality could defeat the justification of Sam’s first-order belief. Wedgwood’s view thus denies the Non-Akrasia Constraint.

2.4 Weatherson and Coates

---

10 Wedgwood [2011], p. 21.
11 Ibid.
12 Ibid.
Finally, Brian Weatherson and Allen Coates each directly address the question of when it could be rational to believe a conjunction of the form, “P, but my evidence doesn’t support P.” Weatherson writes that “[t]he circumstances that could make it reasonable are ones where the agent has good evidence that [P], but also has misleading evidence about the quality of her evidence. … By hypothesis, each conjunct is well-supported by the evidence.”\textsuperscript{13} Coates writes, in a similar vein: “Like all evidence, second-order evidence can sometimes be misleading. When it is, we can rationally judge that our belief is irrational even though it is in fact rational, and so we can rationally be akratic.”\textsuperscript{14} These authors’ general thought is that whenever higher-order evidence is misleading, our first-order attitudes should come apart from our attitudes about what we should believe. So both authors take cases like Sleepy Detective to be counterexamples to the Non-Akrasia Constraint.

To defend this sort of verdict, both authors discuss analogies with moral cases. Weatherson, for example, asks us to imagine the student of a Kantian professor who finds herself facing a murderer at the door, wondering whether to lie about the whereabouts of her roommate. Since her professor’s arguments were so sophisticated and persuasive, Weatherson argues, the student should \textit{believe} that she shouldn’t lie to the murderer; that’s what her evidence supports. Nevertheless, lying to the murderer \textit{is} what she should \textit{do}. Weatherson then suggests that the same kind of situation can occur in epistemology: “sometimes what we should believe is different from what we should believe.”\textsuperscript{15}

The purely epistemic cases that are most analogous to Weatherson’s moral case are those like Sleepy Detective: situations in which an agent responds correctly to some first-order evidence, and then receives evidence that she has not responded correctly to her first-order evidence. In the course of his paper, Weatherson considers a number of

\textsuperscript{13} Weatherson [ms], p. 15. Additionally, in his [2006], Weatherson tentatively endorses the thought that “[i]t is possible for S to have a justified but false belief that her belief in p is justified.” (p. 22) If this is right, a natural consequence is that it could be rational to believe something like, “I should believe P”, while simultaneously disbelieving or suspending judgment on P.

\textsuperscript{14} Coates [2012], p. 122.

\textsuperscript{15} Ibid., p. 12-13. Coates ([2012], p. 116) discusses the story of Huck Finn, who saves Jim, a runaway slave, while believing that he ought to turn him in to his “owner”. The Huck Finn case adds an additional complication: Huck’s belief about what he ought to do is not, according to Coates, epistemically justified. Weatherson’s case is somewhat more straightforward, provided that his account of the Kantian student’s epistemic and moral justification is right.
cases that fit this pattern. He writes,

[T]here are facts about which hypotheses are supported by which pieces of evidence, and ... rational agents do well when they respond to these epistemic facts. ... [T]hese facts retain their normative significance even if the agent has reason to believe that she’s made a mistake in following them. That is, if an agent’s judgment conforms to the correct norms of judgment, then even if she has evidence that she is not good at judging, she should stick to her judgment.

Coates considers similar cases, and concludes that in such situations “one’s total evidence can support both the belief that \( p \) and the judgment that the belief that \( p \) is irrational.” 16

So on these authors’ views, agents in cases like Sleepy Detective are rationally required to be epistemically akratic. 17

2.5 Extreme Level-Splitting

Williamson, Lasonen-Aarnio, Wedgwood, Weatherson and Coates are thus all committed to the following claim: in cases like Sleepy Detective, we can be rationally required to be epistemically akratic. It is clear that these authors’ views will support the verdict that one can be rationally highly confident in \( P \), and also rationally required to be highly confident that one’s evidence doesn’t support as high a level of confidence in \( P \) as one actually has. And it is plausible that these authors’ views will also support the verdict that in more extreme Sleepy Detective-style cases, one can be rationally highly confident in \( P \) and rationally highly confident that one’s evidence supports very low confidence in \( P \).

For the remainder of the paper, I will focus on the possibility that extreme kinds of mismatch can be rational in paradigm cases of misleading higher-order evidence. The question, then, is whether this extreme view can be right: whether in cases like Sleepy Detective, it can be rational to have high confidence in both “\( P \)” and “my evidence doesn’t support \( P \)”, “my evidence supports low confidence in \( P \)”, or “my evidence supports \( \neg P \)”. One might hold that these more extreme states are always irrational, but

---

16 Coates [2012], p. 113.
17 Later in his [ms], Weatherson considers a weakened version of Level-Splitting, whereby HOE would fail to affect Sam’s first-order beliefs only in cases where the first-order belief is the result of a “basic inference”, and where HOE is “relatively weak”. (Weatherson [ms], p. 12.) “Relatively weak” higher-order evidence seems to mean evidence that reduces one to a precise state of neutrality over whether one’s evidence supports \( P \). I see no good explanation for why higher-order evidence should function in this strange way under only these rare, knife-edge conditions. Furthermore, if Weatherson’s moral analogy works in the way it’s supposed to, it seems to support the rationality of more strongly akratic epistemic states. I discuss this further in the next subsection.
maintain that some more moderate mismatches can be rational. Allan Hazlett, for instance, defends a view on which one can rationally believe P while suspending judgment about whether one’s belief in P is reasonable – but one cannot rationally believe P while believing that one’s belief in P is unreasonable.\textsuperscript{18} As Hazlett points out, constraints barring the more extreme form do not entail that the moderate form is also unacceptable.

But while it may be possible to accept a more moderate Level-Splitting view, many of the considerations surveyed above clearly support an extreme version. Weatherson’s moral analogy, for example, relies on the idea that one can be morally required to take some action (i.e. lie to the murderer) while being rationally required to believe that one should take a different, incompatible action (i.e. tell the truth). The most direct epistemic analogy would be a case in which one believes P, but also believes that one should believe ~P: a paradigm case, that is, of extreme epistemic akrasia. On Wedgwood’s view, the rationality of an inference is entirely a matter of the particular mental process. If that process can occur even when an agent harbors “serious doubts” about its reliability, a natural consequence is that an agent’s belief in P can be rational despite her believing that she should not believe P. If these motivations support any kind of Level-Splitting view, they support one on which extreme epistemic akrasia can be rational. And this consequence is clearly a welcome one, to at least some authors; as we’ve seen, some of the authors quoted above explicitly extend their views to the extreme cases.

If extreme divergence is irrational, we can rule out extreme Level-Splitting views completely. Moreover, looking at this strong kind of divergence is worthwhile even for those looking to support weaker Level-Splitting views: seeing why extreme epistemic akrasia is irrational can shed light on whether a moderate Level-Splitting view could be supported, and if so, what kind of desiderata such a view must meet. For present purposes, then, I will focus on Level-Splitting views on which the full range of epistemically akratic attitudes can be rational in cases like Sleepy Detective.

\section*{3. Immediate problems}

\textsuperscript{18} Hazlett [2012].
If Sam follows Level-Splitting’s advice, he will come to have high confidence in something like, “Lucy is the thief, but my evidence doesn’t support that.” This, by itself, already looks strange; the conjunction of the two statements has a somewhat Moore-paradoxical quality. 19 Although statements of this form are not contradictory — both conjuncts will be true whenever one’s evidence is misleading — there is something odd about believing or asserting them simultaneously.

Since Level-Splitters claim that Sam is rationally required to have high confidence in both statements, they won’t be moved very far by contentions of the conjunction’s intuitive oddness. But it’s worth pointing out that the oddness extends beyond the combination of attitudes itself: the counterintuitive consequences Level-Splitters must accept are more far-reaching than one might have thought. I will go through a few immediate examples here.

First, suppose that Sam, having reacted as Level-Splitting recommends, reflects on his situation. He is highly confident that Lucy is the thief. And he has some first-order evidence, E, which he is highly confident does not support the proposition that Lucy is the thief. What should Sam think about his epistemic state? If Sam takes both “Lucy is the thief” and “my evidence doesn’t support Lucy” as premises, it seems that he can engage in some patently bad reasoning. 20

More specifically, first: what should Sam think about his first-order belief that P? Upon reflection, Sam might naturally wonder how he came to have this particular true belief. Usually, we come to have true beliefs by correctly evaluating our evidence. But Sam doesn’t think that this is what’s happened in his case: he believes that his evidence doesn’t support P. So perhaps he should just think that he got lucky. “Despite the odds,” he should think, “I’ve come to the truth about P!”

The oddness of Sam’s attitude toward his belief that P might look even stranger if we imagine another situation in which Sam first receives some higher-order evidence to the effect that he is going to be unreliable in assessing some evidence, E, in the future. Sam should then become highly confident that whatever judgment he makes at that future time will be false; after all, that’s usually what happens when we make irrational

19 See Smithies [2012] for further discussion of this point.
20 A Level-Splitter might object to this point by denying that Sam is rationally permitted to reason from both of these propositions. I will address this objection in section 5.
judgments. But when the future time comes and Sam assesses E, finding himself inclined to judge that P on its basis, what should he think? According to the Level-Splitting view, he should be highly confident that P. “I thought I was going to judge falsely,” Sam might say to himself, “but I must have I lucked out! I judged that P, and P is true.”

Second, what should Sam think about his first-order evidence, E? Well, he is rationally highly confident that E bears on P; E is all of his evidence about the jewel thief case, and he knows that his evidence will strongly support the guilt of one of the suspects. This means that it is rational, given E, to have high confidence in the guilt of one suspect and low confidence in each of the others: for each suspect, Sam’s rational confidence that she is guilty, on E, should either be quite high or quite low. On the basis of his higher-order evidence, Sam is rationally highly confident that E doesn’t support P – that Lucy is the thief. So, Sam should think, most likely he should have quite low confidence in P. But because of his first-order evidence, Sam’s confidence in P is quite high. It seems that he can then put the pieces together and reason in the following way: “P is true. But all of my evidence relevant to P does not support it. It supports low confidence in a true proposition, P, and therefore high confidence in a false proposition, ~P. So, E is misleading.”

Is this a legitimate way to draw the conclusion that one’s evidence is misleading? To better see what’s wrong with Sam’s reasoning, consider an analogous case: Sam is driving in to work one morning, and plans to evaluate some evidence, E, when he gets there. He hasn’t seen E yet, but as he is driving, he rationally thinks to himself: “Whatever E is, it is most likely not misleading – after all, good evidence usually supports the truth.” When he walks in the door, Alex tells him (falsely) that she has spiked his coffee with a reason-distorting serum. “That’s too bad,” Sam thinks. “Now I won’t be able to evaluate my evidence correctly; I’m unlikely to find out the truth about the jewel thief.” Then Sam opens the door to the evidence room, and sees E before him. E, in fact, supports P. Just as in Sleepy Detective, Sam evaluates E rationally, concludes that P, and takes his higher-order evidence to indicate that E does not support P. “Aha!” he exclaims, “P is true! But this evidence doesn’t support P; it’s misleading. That’s too bad.” Then Sam thinks another minute, and smiles. “Actually, it’s a good thing Alex
spiked my coffee. If I had evaluated my evidence correctly, I would have ended up with a false belief.”

Sam’s reasoning here does not look rational: he should not be able to conclude that $E$ is misleading in this way. But we should be careful before condemning his argument: there are circumstances in which it is rational to conclude that you have misleading evidence. For example, it can be rational to believe that a particular subset of your evidence is misleading with respect to a particular proposition. This might happen when you receive new, more conclusive evidence, showing that your original evidence pointed in the wrong direction. So you can rationally believe $P$, but also rationally believe that part of your evidence supports $\neg P$. In this type of situation, we can tell a sensible story about why your belief state is stable. By basing your belief on your total evidence, you have avoided being misled. So although you realize that you have misleading evidence, there is no pressure to revise your beliefs.

It can even be rational, in some cases, to conclude that your total evidence is misleading. For example, in a lottery case, you can be rationally highly confident that one of the propositions of the form “Ticket $n$ will lose” is false, though you are also rationally highly confident in each of these propositions. Here, we can similarly tell a story about why your belief state is stable. Since you cannot identify which ticket proposition is false, you cannot identify the particular way in which your evidence is misleading. There is no particular other belief state that you think would be more accurate than your own belief state; therefore, there is no way for you to avoid being misled. So, it seems, you should not revise your beliefs.

But Sam’s case is not like either of these: it seems that Sam can avoid being misled. He can point to a particular belief of his that is, he thinks, unsupported by his total evidence. He may even be able to identify a particular belief state that, he thinks, he should rationally adopt. (For example: supposing that Alex told him that his reasoning made him no more reliable than chance, Sam might think that the rational thing for him to do is to spread his credence evenly over the various candidates.) Unlike in the lottery case, for example, there is something in particular that Sam thinks he should believe – but he does not believe it.

---

21 Thanks to an anonymous *Nous* referee for suggesting that I consider this type of case.
If Level-Splitting is right, and extreme cases of epistemic akrasia can be rational in Sleepy Detective, there is nothing wrong with Sam’s concluding that his evidence is misleading in this way. All Sam needed to get this reasoning started was high confidence in both “P” and “my evidence doesn’t support P”, plus the setup of the case: that E bears strongly on P. According to Level-Splitting, being in this kind of akratic state in situations like Sam’s can be perfectly rational. But there is something wrong with Sam’s concluding that his evidence is misleading in this case. This suggests that there is something wrong with Level-Splitting: specifically, there is something wrong with its denial of the Non-Akrasia Constraint.

So far we have seen two immediate counterintuitive consequences of Sam’s reasoning from his akratic state in Sleepy Detective. Another way to bring out the oddness of Sam’s belief state is to consider what it would look like to act on his beliefs; the irrationality of an agent’s beliefs is often brought into sharper focus if we consider what happens when she puts her money where her mouth is. Perhaps if Sam is very confident that P is true, he will be willing to make a wager:

Sam: I’d bet that it’s Lucy. I’ll give you 9:1 odds.
Alex: But you were so sleepy when you were working last night! How can you be so sure that the evidence supports her guilt?
Sam: Oh, I’m not. Since you told me that I’m so bad at evaluating evidence when I’m tired, I doubt the evidence supports Lucy’s guilt much at all. If I were to bet on what the evidence supported, I might give you 1:9 odds that it’s Lucy, but certainly not 9:1.
Alex: So why are you offering 9:1 odds?
Sam: Well, I really shouldn’t be offering such strong odds. I shouldn’t be so confident that she’s the thief: the evidence isn’t in my favor. But on the other hand, she is the thief! That’s what we’re betting on, right?

Sam’s behavior here illustrates the irrationality of believing as he does. It seems patently irrational to treat a bet about P and a bet about whether one’s evidence supports P as completely separate. But if Level-Splitting is right, this is just how a rational person will behave in Sam’s situation.\textsuperscript{22}

\textsuperscript{22} Coates discusses a similar betting argument, and suggests that it can be avoided by holding that agents in Sam’s situation should not act on their akratic belief states. He writes, “it is a [practically] rational policy not to make bets one believes are irrational, since acting on this policy will likely maximize expected value in the long run”. (Coates [2012], p. 121.) In other words, instead of following the rule “take bets that
And if Sam can act on his akratic beliefs, his odd-looking behavior won’t be restricted to betting. Should Sam tell newspaper reporters that Lucy is the thief? Recommend that the cops raid her apartment? Send her to jail? If he is rationally highly confident that she is guilty, it seems that all of these actions could be warranted. But if Sam is asked to justify or explain his behavior, he will be at a loss: as far as he thinks, his evidence doesn’t support any of it. In fact, he might think that he shouldn’t be acting in these ways.  

Level-Splitting is thus committed to counterintuitive results beyond the initial strangeness of believing both “P” and “my evidence doesn’t support P” in Sleepy Detective. It also seems to license patently bad reasoning and irrational action.

In the next section, I will look at some further drawbacks to the Level-Splitting view – among them, a different sort of bad reasoning – resulting from attempts to reconcile a Level-Splitting verdict in Sleepy Detective with intuitive verdicts in related cases.

4. Problems with extending Level-Splitting

So far we have seen that Level-Splitting gives a certain verdict in cases like Sleepy Detective: Level-Splitting says that in such cases, one’s confidence in P should be the level of confidence supported by one’s first-order evidence alone, and one’s confidence in “my evidence doesn’t support P” should be the level of confidence supported by one’s higher-order evidence alone. (Recall Weatherson’s claim: “by hypothesis, both conjuncts are well-supported by the evidence.”) But what should Level-Splitting say about other cases? Should epistemic levels always operate separately? Or should higher-order evidence rationally affect one’s first-order attitudes under some circumstances (though maximize expected value”, we should follow something like “take bets that you believe will maximize expected value.” Coates acknowledges that following this second rule will lead one to not maximize expected value in some particular cases; so, we would in fact do better by following a policy of maximizing expected value. In order to take Coates’ response on board, we must either deny that it is rational to maximize expected value, or else posit two types of rational evaluation. For example, there might be one norm that tells us to follow the rules that would be best if followed, and one norm that tells us to follow the rules that would be best if we tried to follow them. (This second proposal is perhaps comparable to Lasonen-Aarnio’s epistemic norm of “reasonableness”; see fn 8.) Coates’ response deserves further discussion, but I will not pursue the suggestion here.

23 Again, a Level-Splitter might object to this point by denying certain bridge principles between rational degrees of belief and action. I will return to this objection in section 5.
not those in Sleepy Detective)? In this section I’ll look at some possible ways one might extend the Level-Splitting view, and suggest that none of the options looks good.

4.1: No interaction, ever

A very strong version of Level-Splitting, on which higher-order evidence is never relevant to first-order attitudes, would be a non-starter. To see why, suppose that Sam and Alex are working on another case. All of the relevant evidence is down at the station; Sam is on his way in to work, and plans to evaluate the evidence when he gets there in order to determine which of the suspects is guilty. But Alex gets to the station first.

Before Sam has seen any of the relevant evidence, Alex calls him up and tells him:

Alex: I’ve figured it out. All of the evidence points to Veronica’s guilt!
Sam: Well, I don’t care what the evidence points to. Let me see the evidence, so I can figure out who did it.
Alex: Veronica did it.
Sam: You could have said so in the first place! I thought you were just telling me what the evidence supported.

Sam’s response here is irrational. When Alex tells him what the evidence supports, Sam should become confident that Veronica is guilty. But a strong, no-interaction-ever Level-Splitting view would not allow this. So Level-Splitters should not – and, I think, would not want to – adopt this extreme view.

4.2: Interaction only in the absence of relevant first-order evidence

A better Level-Splitting view will need to allow higher-order evidence to rationally affect Sam’s first-order beliefs in situations like the one described above, while denying that higher-order evidence rationally affects first-order beliefs in situations like Sleepy Detective. What is the difference between the two types of cases? One attractive hypothesis is that in Sleepy Detective, Sam already has the relevant first-order evidence. But in the case described in section 4.1, he does not. And this distinction is, plausibly, a significant one. Higher-order evidence is evidence about what a body of evidence supports. A Level-Splitter might argue that when higher-order evidence just bears on your evidence – that is, a body of evidence that you already have – it has no rational
effect. But when higher-order evidence bears on a body of evidence that you don't have, it is rationally significant.

If a Level-Splitter decides to go this way, one thing she might have in mind is something like this principle:

**Proxy:** Higher-order evidence is relevant to first-order beliefs only insofar as it serves as a proxy for first-order evidence.

Proxy isn't explicitly endorsed by any of the people quoted in Section 2; their discussions of higher-order evidence are generally restricted to cases where one's higher-order evidence bears on first-order evidence that one already has, so the issue doesn't arise. But the idea behind Proxy is a fairly intuitive one, and could easily be adopted by any of the people quoted above. So before moving on, I will briefly look at Proxy, how it might motivate Level-Splitting, and how it might run into problems; this will give us a better idea of how higher-order evidence should work.

First, let's see how Proxy might deliver the Level-Splitting verdict in the case above, where Alex tells Sam about evidence that he hasn't seen yet. When Alex tells Sam about the evidence she has seen, she is telling him something about what kind of evidence is out there. Since Sam knows a lot about detective work, it's possible that Alex's testimony tells him something about the specific first-order evidence that she saw. So when Alex tells Sam that all of the evidence points to Veronica's guilt, perhaps Sam should increase his confidence in the long disjunction of first-order evidential propositions that would point to Veronica. (That is: that Alex has a gun with Veronica's fingerprints on it, or a letter in which Veronica describes her plans to an accomplice, or...) Since the truth of any of these disjuncts would itself support P, in becoming confident of the disjunction, Sam should also become more confident of P.

Second, Proxy could deliver the Level-Splitting verdict in cases where one learns that there is evidence for P, without identifying what that evidence is. For example, suppose that Sam, who knows nothing about physics, befriends a physicist. The physicist tells him that all of the evidence in her field strongly supports the proposition that the universe is expanding. Sam has no idea what such evidence would look like. (In fact, we can imagine that Sam knows so little about physics that for any candidate evidential proposition he can think of, he doesn't know whether it would count in favor of or
against the first-order proposition in question.) Surely, in this case, Sam should become more confident that the universe is expanding. And Proxy can predict this.

So if a Level-Splitter endorses Proxy, she can maintain that higher-order evidence is relevant to first-order beliefs in some cases. She can also explain why, on her view, higher-order evidence is not relevant to Sam’s first-order belief in Sleepy Detective. In Sleepy Detective, Sam already had all of the relevant first-order evidence. So it is implausible that his higher-order evidence could serve as a proxy for first-order evidence in either of the ways mentioned above. Sam’s higher-order evidence should not increase his confidence in a long disjunction of first-order evidential propositions; he already knows which photos, letters, etc. compose the first-order evidence. And Sam’s higher-order evidence should not increase his confidence that there is some first-order evidence at the police station that supports Veronica’s guilt. That’s because Sam already knows exactly what first-order evidence there is.

If higher-order evidence only affects first-order beliefs by Proxy, it will have no effect in cases where one already has the first-order evidence. In those cases, higher-order evidence will be screened off.

These observations suggest that a more plausible version of Level-Splitting, motivated by Proxy, should endorse something like the following prediction:

Higher-order evidence should rationally affect one’s first-order beliefs only insofar as one does not also possess the relevant first-order evidence.

When one has both first-order and higher-order evidence, a Level-Splitter can claim that first-order evidence should affect one’s first-order beliefs, and higher-order evidence should affect one’s beliefs about what one’s evidence supports. But she is also free to argue that, in these cases, higher-order evidence should not affect one’s first-order beliefs. Proxy, then, seems to give Level-Splitters a tidy and intuitive way to distinguish between cases where higher-order evidence should rationally affect one’s first-order beliefs and cases in which it should not.

But there is good reason to doubt that Proxy gives the whole story about how higher-order evidence works. Compare the following two cases:

24 Weatherson considers this position, which he calls “ESJ” (for “Evidence Screens Judgment”) in his [ms]. Though he doesn’t explicitly endorse the position, much of what he says is consistent with the thought that, at least as far as first-order beliefs go, first-order evidence screens off higher-order evidence.
**Case 1:** Sam is working to find a serial killer. He has his evidence, $E$, set out in front of him, and he has carefully looked through all of it. $E$ supports the proposition $M$: that Marie is the killer. Sam has not yet finished his analysis of $E$ and has not reached a conclusion about what $E$ supports. Nevertheless, Sam just happens to believe – for no particular reason – that Marie is the killer.

**Case 2:** Sam and Alex are working together to find a serial killer. They have their evidence, $E$, set out in front of them, and they have both carefully looked through all of it. $E$ supports the proposition $M$: that Marie is the killer. Sam has not yet finished his analysis of $E$ and has not reached a conclusion about what $E$ supports. But Alex finishes her analysis first, and exclaims, “I've got it! The evidence points to Marie.” On the basis of Alex’s testimony, Sam comes to believe that Marie is the killer.

These two cases have some things in common. In both cases, Sam’s belief that $M$ is supported by his first-order evidence, $E$. So there may be a sense in which Sam is required to believe $M$ in both cases. But there also seems to be an important asymmetry between Case 1 and Case 2. In particular, Sam’s belief that $M$ seems much better justified in Case 2 than in Case 1. One way to put the point is like this: while Sam’s belief that $M$ is propositionally justified in both cases, it is only doxastically justified in Case 2.25

What could explain this asymmetry? Doxastic justification requires basing one’s beliefs on evidentially relevant factors. So if Sam’s belief in $M$ is doxastically justified in Case 2, it must be because the basis for his belief in $M$ – his higher-order evidence from Alex – is evidentially relevant to $M$. But Proxy does not predict this. After all, in Case 2, as in Sleepy Detective, Sam already knows what his first-order evidence is. Proxy says that when one has the relevant first-order evidence, higher-order evidence is not at all evidentially relevant to one’s first-order beliefs. So if we want to capture the asymmetry between Case 1 and Case 2, we must allow that higher-order evidence can rationally affect first-order beliefs in ways other than by Proxy. A view on which Proxy is the only connection between epistemic levels will be missing something.

Perhaps Proxy points to one way in which higher-order evidence rationally affects first-order beliefs. But Case 2 suggests that this cannot be the only way.

Rather, it seems that higher-order evidence must also affect first-order beliefs by

---

25 One might object that Sam’s belief that $M$ in Case 2 is not fully doxastically justified, since it is not based on his total relevant evidence. But even if we grant this, there still seems to be an asymmetry between the two cases: Sam’s belief that $M$ seems to be somewhat doxastically justified in Case 2, and is not at all doxastically justified in Case 1.
bearing on evidential support relations: it is evidence about what some body of
evidence supports.

Case 2, as a counterexample to Proxy, also spells trouble for Level-Splitting
more generally. That’s because any Level-Splitter who tries to accommodate this
counterexample will have a hard time maintaining her verdict in Sleepy Detective. In
Case 2, it seemed plausible that Sam’s higher-order evidence should rationally bear
on his first-order beliefs by way of bearing on evidential support relations. In Sleepy
Detective, Sam’s higher-order evidence seems to work in the same way, by bearing
on the evidential support relation between E and P. So it’s natural to think that Sam’s
higher-order evidence should have parallel effects in the two cases: if it rationally
affects Sam’s first-order beliefs in Case 2, it should do so in Sleepy Detective. Level-
Splitters are therefore left with two options. They may either deny that there is any
asymmetry between Case 1 and Case 2, or find a significant disanalogy between Case
2 and Sleepy Detective to justify treating the two cases differently. Neither option
looks promising.

4.3: Interaction in the long run
In this section we have so far looked at two questions about how the Level-Splitting view
should be extended. We have seen that a plausible Level-Splitting view should allow
higher-order evidence to affect first-order attitudes in some cases, but it turned out that
one particularly tidy way of dividing these cases is unlikely to work.

A final question to consider is how Level-Splitting should deal with cases like
Sleepy Detective if they happen again and again. Answering this question will be
somewhat complicated, but looking at how a Level-Splitter might deal with a long string
of cases will help further bring out what is wrong with Level-Splitting views.

Recall that Level-Splitting says that in cases like Sleepy Detective, one should
become confident in P to the level supported by one’s first-order evidence alone, and one
should become confident in “My evidence doesn’t support P” to the level supported by
one’s higher-order evidence alone. So in a single case like Sleepy Detective, Sam should
become highly confident in each of those propositions.
What if Sam finds himself in cases like Sleepy Detective many times in a row? (We can imagine this happening in different ways: perhaps on the same night, Sam evaluated dozens of cases before talking to Alex on the phone; alternatively, several cases like this could occur sequentially.) If this happens, Sam will acquire what will seem to him like a surprisingly good track record: he will have high confidence in lots of true propositions, despite being quite unreliable at rationally assessing his evidence about those propositions. So Sam might reason along the following lines: “I believe P, Q, and R, and furthermore, P, Q, and R are true. Yet I formed my beliefs that P, Q, and R irrationally; I was rationally impaired (drunk, sleepy, etc.) when I did the calculations. It’s highly unlikely that I would get all of these answers right, given that I wasn’t rational in arriving at them.”

According to Level-Splitting, what should Sam think has happened? I will consider two possible ways a Level-Splitter could answer this question, and suggest that either way raises new challenges for the view.

The first possible answer is that, even after the end of a long series of cases, Sam should remain epistemically akratic. That is, he should maintain his high confidence in P, Q, R, etc., and he should maintain his high confidence that his evidence doesn’t support P, Q, R, etc. After a very long string of cases, how should Sam explain his success? Perhaps he should think he is extremely lucky. Strange things sometimes happen, and Sam could think that his long string of epistemic successes is one of them. Or perhaps Sam could increase confidence in some other hypothesis – one that would explain why he keeps getting things right, despite his unreliability at evaluating his evidence. If Sam goes this route, perhaps he should think that he has supernatural powers, or that a guardian angel is guiding him to the truth.

But none of these hypotheses seems like the best explanation for Sam’s situation. After an apparently fair coin lands heads a hundred times in a row, you should not end up with high confidence that it’s all a lucky coincidence, or that there is a heads-loving coin fairy. Instead, you should reduce your confidence that the coin was ever fair in the first place. Analogously, the best explanation for Sam’s situation is that he was rational all along: as he gets things right again and again, Sam should wonder if maybe he is better at evaluating his evidence than Alex suggested, and perhaps his evidence really supported
his beliefs all along. So if Level-Splitters say that Sam should remain akratic at the end of a long sequence like this, they render him unable to rationally draw the conclusion that best explains his situation.

A second possibility for Level-Splitters is to say that, by the end of a long sequence of cases, Sam should no longer be epistemically akratic. The argument might go something like this. As Sam acquires a track record of true beliefs about $P$, $Q$, $R$, etc., he might take this track record to be evidence for his own reliability. The best explanation for reliability in reaching true beliefs is reliability in assessing one’s evidence. So after a long enough string of successes, Sam should conclude that he was never seriously rationally impaired in the first place. That is, he should conclude that his evidence did support $P$, $Q$, and $R$. So he should not be akratic.

In adopting this strategy, Level-Splitters can claim that Sam should, in fact, draw the conclusion that best explains his situation. But the process by which Sam can come to this conclusion brings problems of its own. Sam’s method of concluding that he was rational is an instance of bootstrapping – assessing the reliability of a belief process by using the process itself, without employing any independent checks. Bootstrapping in any context is notoriously suspicious.\(^{26}\) In fact, Christensen raises the possibility of bootstrapping as an argument against the rationality of being epistemically akratic in a certain way: being certain of some logical theorem, but also doubting that one is perfectly rational.\(^{27}\) If Level-Splitters say that Sam should bootstrap, their first challenge is to provide an explanation for why bootstrapping is an acceptable way to reason.

Of course, there are also bootstrapping apologists, and a Level-Splitter could always adopt that strategy too. For example, she might point out that, on her view, Sam can only rationally bootstrap if he has in fact evaluated his first-order evidence correctly in a large number of cases – so he will at least be prevented from bootstrapping to a false conclusion.\(^{28}\)

\(^{26}\) See, e.g., Vogel [2000] and Cohen [2002].

\(^{27}\) See Christensen [2007 b], p. 18. Christensen’s discussion combines the possibility of bootstrapping with the possibility of extreme luck; if what I say here is right, however, Level-Splitters should see these possibilities as very different. See also White [2009]. Elga [2007] makes a similar argument in the context of peer disagreement.

\(^{28}\) Kornblith [2009] defends reliabilism on roughly these grounds. Kelly ([2007], esp. section 5.4) makes a similar argument defending his “Total Evidence” view of peer disagreement against Elga [2007]’s objection that Kelly’s view sanctions bootstrapping.
But even if this response is satisfactory, Level-Splitters must meet a second, and more serious, challenge in defending the claim that Sam should bootstrap. The most natural way to get the bootstrapping reasoning off the ground is to note that truly believing P when one’s evidence does not support it is unlikely or unexplained. If we need an explanation for being in such a state again and again, presumably it is because there is a kind of rational tension involved in being in this state without an explanation. Otherwise it would be perfectly rational to remain akratic even in a large number of cases. The thought behind this reasoning seems to rely on level-bridging: the idea that there is some kind of rational connection between epistemic levels. But level-bridging appears to be inconsistent with Level-Splitting: if there is a rational connection between epistemic levels, why doesn’t higher-order evidence rationally affect our first-order beliefs in cases like Sleepy Detective?

Any level-connection that would explain why bootstrapping is rational here threatens to undermine the Level-Splitting verdict in a single case. And the reasoning that explains the Level-Splitting verdict in a single case (for example, Proxy) seems unable to help justify bootstrapping. So if Level-Splitters want to say that Sam should no longer be akratic after a large number of cases like Sleepy Detective, they face the challenge of reconciling these two apparently incompatible claims: that in a single case, epistemic levels operate separately, but in the long run, they do not.

The bootstrapping argument is therefore one more instance of a general challenge for Level-Splitting. A Level-Splitter must maintain her verdict in a single case like Sleepy Detective, but should also allow that in certain similar situations — situations where Sam does not have the first-order evidence, for instance, and situations where the conditions in Sleepy Detective come about again and again — there is a rational connection between one’s first-order beliefs and one’s beliefs about what one’s evidence supports. If Level-Splitters cannot tell a story about what kind of level-connection might support both judgments, they face a dilemma: they must either embrace absurdly strong Level-Splitting verdicts in a variety of situations, or give up their verdict in Sleepy Detective.

5. Bridge principles, reasoning, and action
I've argued against having high confidence in “P, but my evidence doesn’t support P” on
the grounds that being in an akratic state like this, in cases like Sleepy Detective, licenses
bad reasoning and irrational action. In order for these arguments to go through, I assumed
that both “P” and “my evidence doesn’t support P” were available to an akratic agent like
Sam for the purposes of rational reasoning and action. These assumptions could be
thought of as bridge principles, perhaps looking something like this:

**Reasoning Bridge Principle:** It is rational to reason from any proposition, P, just
in case one is rationally highly confident in P.\(^{29}\)

**Action Bridge Principle:** It is rational to act on any proposition, P, just in case
acting on P maximizes expected value.

A Level-Splitter might then object: why hold these bridge principles? Without them, we
need not say that epistemic akrasia licenses bad reasoning or irrational action. So in order
to salvage the Level-Splitting view, the objection goes, we can just reject the bridge
principles.

There are several ways one might fill in the details of this objection. For instance,
one might argue that it is irrational to reason (or act) using beliefs that do not amount to
knowledge (in this case, “my evidence doesn’t support P”).\(^{30}\) Alternatively, one might
argue that it is irrational to reason (or bet) using beliefs that one believes to be unjustified
(in this case, “P”).\(^{31}\) Either way, an epistemically akratic agent like Sam wouldn’t be able
to rationally reason or act on both “P” and “my evidence doesn’t support P”. I will
respond to this objection here.

First, we should note that giving up the Reasoning and Action Bridge Principles is
not an obvious or easy move to make. For one, the Level-Splitting view and the
motivations that count in its favor are at least prima facie independent of these bridge
principles. So denying the bridge principles is a substantial further claim, and would
require additional argument.

\(^{29}\) Of course, someone defending this principle should also require that one’s reasoning preserve coherence,
respect one’s total evidence, and the like. Spelling out these details might best be thought of as giving
norms of good reasoning; the bridge principle I discuss here is just supposed to determine which
propositions one can rationally use in one’s reasoning in the first place.

\(^{30}\) Williamson, for example, might want to go this way. Since knowledge is the norm of both assertion and
belief, on his view, it is plausible that the view might also prohibit one from reasoning from premises that
are not known.

\(^{31}\) Coates’ discussion of the betting argument employs a version of this strategy. See fn 22.
And there is reason to think that this additional argument will be a tough sell: the bridge principles (or principles like them, at any rate) have considerable intuitive appeal. It's plausible that rational belief is closely connected with rational action, and that rational patterns of reasoning are determined by what one’s evidence supports. And in order to block all of the apparently problematic consequences of epistemic akrasia, Level-Splitters will need to place significant restrictions on the bridge principles. (In order to say, for example, that Sam should not tell the police to arrest Lucy, Level-Splitters will have to go farther than denying that one can rationally place bets or make esoteric arguments.) These connections are, plausibly, part of what makes epistemic rationality worth theorizing about in the first place: looking at what we should believe helps us figure out what we should do. Those who are interested in epistemic rationality for these reasons should be especially hesitant to adopt a view on which the bridge principles do not hold.

Second, even if we give up the Reasoning and Action Bridge Principles, the arguments that appealed to these principles still have some bearing. To see why, let’s just focus on Sam’s argument that his evidence is misleading. The problem there seemed to be that Level-Splitting allowed Sam to irrationally draw a certain conclusion. If Level-Splitting denies the Reasoning Bridge Principle, the objection suggests, this problem would be avoided. But even if Sam cannot rationally draw the conclusion – for example, because he is prohibited from consciously going through the relevant argument, or because the justification for the argument’s premises (“P” and “E supports low confidence in P”) is not “transmitted” to the conclusion – the argument still shows that something is wrong with Level-Splitting. “P” and “E supports low confidence in P”, together, entail the conclusion that E is misleading with respect to P. So regardless of whether or not Sam can use this conclusion to argue, act, or form new beliefs, his evidence supports it. That seems troubling enough. If it is implausible that Sam should believe that E is misleading, it is also – and for the same reasons – implausible that his evidence should support it.

The same goes for the other examples of bad reasoning or irrational action. The examples of bad reasoning show that, if Level-Splitting is right, in cases like Sleepy Detective our evidence supports odd and counterintuitive conclusions beyond “P, but my
evidence doesn’t support P”. The examples of irrational action bring out what is intuitively problematic about representing the world in the way that, according to Level-Splitting, is supported by our evidence. Even without the bridge principles, both kinds of argument give us reason to worry about the evidential support relations endorsed by Level-Splitting. And since Level-Splitting is a view about rational belief, a reason to worry about Level-Splitting’s evidential support relations just is a reason to worry about the view itself.

6. Possible exceptions?
Sleepy Detective is a paradigmatic example of the type of case usually discussed in the higher-order evidence debate, and so far we have seen several problems with the view that epistemic akrasia is rational in that case. But recent literature has suggested that under certain special conditions, higher-order evidence might work differently from how it does in cases like Sleepy Detective. And some authors – even some authors who defend Conciliatory views about cases like Sleepy Detective – have argued that under those special conditions, it is rational to be epistemically akratic. What does this mean for the Non-Akrasia Constraint, and the arguments we have seen so far in its support?

I will close by considering an example of this type of special case, which I’ll call “Dartboard”. As we will see, epistemic akrasia seems rational in Dartboard only if we make the (controversial) assumption that we can be rationally uncertain about what our evidence is. One might wish to deny that assumption for independent reasons, in which case Dartboard will not seem like much of a worry. But for those who find the assumption plausible, admitting epistemic akrasia in Dartboard looks like serious cause for concern. To answer this concern, it will be worth looking at the relationship between Dartboard and Sleepy Detective.

In this section I will argue that accepting a pro-akrasia verdict in Dartboard does not commit us to Level-Splitting. On the contrary, examining why epistemic akrasia might be rational in this type of special case will help us see why Level-Splitting is wrong. I will also suggest that, surprisingly enough, those who accept pro-akrasia

32 See, e.g., Christensen [2010] and Elga [forthcoming].
verdicts in cases like Dartboard can still argue against Level-Splitting by appealing to some of the main anti-akrasia arguments that I have offered here.33

I will focus on the following scenario, adapted from Williamson;34 it is similar to Williamson’s puzzle of the unmarked clock:35

**Dartboard:** You have a large, blank dartboard. When you throw a dart at the board, it can only land at grid points, which are spaced one inch apart along the horizontal and vertical axes. (It can only land at grid points because the dartboard is magnetic, and it’s only magnetized at those points.) Although you are pretty good at picking out where the dart has landed, you are rationally highly confident that your discrimination is not perfect: in particular, you are confident that when you judge where the dart has landed, you might mistake its position for one of the points an inch away (i.e. directly above, below, to the left, or to the right). You are also confident that, wherever the dart lands, you will know that it has not landed at any point farther away than one of those four. You throw a dart, and it lands on a point somewhere close to the middle of the board.

We can illustrate Dartboard like this (suppose the board is unmarked and much larger, so that it is not obvious where the dart is):

```
1  2  3  4  5
1  
2  
3  
4  
5  
```

Let’s assume the dart landed at $<3,3>$. Given that the dart landed at $<3,3>$, what should you believe about where the dart landed?

In this case, many authors find it intuitively plausible that, for any particular point, you should not be certain that that point is where the dart landed. After all, it’s a

---

33 Special thanks to Bernhard Salow for helpful discussion on this topic.
34 See Williamson [2000], p. 229. Williamson presents the case abstractly; Dartboard is a slightly more realistic version with the same formal properties.
35 The original clock case is from Williamson [notes a] and [notes b], and is also discussed in Christensen [2010] and Elga [forthcoming]. The only difference between Dartboard and the clock case is that in Dartboard, the evidence varies along two dimensions rather than one. I discuss Dartboard rather than the clock because, in a two-dimensional case, it is more plausible that there can be a large divergence between first-order and higher-order attitudes: that is, that it can be rational to be highly confident of both “P” and “my evidence supports ¬P”, rather than the more moderate-sounding “P” and “my evidence supports slightly lower confidence in P than I actually have”. That is, Dartboard seems to motivate “extreme” epistemic akrasia in the sense I discuss in section 2.5.
large blank board, and you have good reason to think that you might confuse adjacent points with one another. Nevertheless, you should be able to narrow down the possibilities significantly: your discrimination may not be perfect, but it's still pretty good. So let's suppose that when the dart lands at \( <3,3> \), you should be highly confident in the proposition that it landed at either \( <3,2>, <2,3>, <3,3>, <4,3>, \) or \( <3,4> \) – so, you can rationally rule out every point except for those five. (If this margin for error seems implausible, expand or shrink your imagined dartboard as necessary.) Williamson agrees with this verdict, and supposes further that your credence should be equally distributed over \( <3,2>, <2,3>, <3,3>, <4,3>, \) and \( <3,4> \). So, for each of those five points, you should have .2 credence that the dart landed at that point.

Now consider the proposition Ring: that the dart landed on one of \( <3,2>, <2,3>, <4,3>, \) or \( <3,4> \). (That is, that it landed on one of the points in the ring around \( <3,3> \).) Your credence in Ring should be .8. But that level of confidence in Ring is only rational if the dart actually landed on \( <3,3> \). If the dart had landed on some other point, the rational credence distribution would be centered on that point instead of \( <3,3> \), and your rational credence in Ring would be lower than .8. In particular, if the dart landed at any of \( <3,2>, <2,3>, <4,3>, \) or \( <3,4> \), it would be rational for you to have .2 credence in Ring. Suppose you have reflected on your situation, and you know what would be rational to believe in these various evidential situations. So you should be .8 confident of Ring, but also .8 confident that your credence in Ring should be .2. This means you should be epistemically akkratic: you should be highly confident of both “Ring” and “my evidence supports ~Ring”.

How should we think about this case, given that we have already seen how bad epistemic akkrasia is in Sleepy Detective? One possibility is to reject Williamson's verdict. Perhaps we could do this by moving to an extreme view on which we should never be rationally uncertain of what our evidence is or what it supports; this would rule

---

You might object that this extreme result only comes about because we assumed that your credence should be distributed evenly over the five possibilities. In a more realistic case we should want to give \( <3,3> \), the point where the dart actually landed, at least a bit more weight. It's true that the divergence between your credence in Ring and your credence that your evidence supports ~Ring will get smaller and smaller as we put more weight on \( <3,3> \), but the general problem will not go away so easily. In a three- or four-dimensional case, putting a lot of credal weight on the actual case becomes less plausible. So it would be hard to admit that Williamson's description of Dartboard-style cases is ever right without also admitting that there are some cases in which extreme epistemic akkrasia is rational.
out all kinds of epistemic akrasia as rational, including that motivated by Dartboard. But if our main interest is to preserve an anti-akrasia verdict in Sleepy Detective, such extreme measures are not necessary. For the remainder of this section, I will adopt the assumption that Williamson’s verdict in Dartboard-style cases is right, and look at how this verdict squares with an anti-akrasia view about Sleepy Detective.

One difference we should notice right off the bat is that the akratic state required in Dartboard is somewhat different from the akratic state that Level-Splitters require in Sleepy Detective. In Sleepy Detective, Level-Splitters say that you can rationally (a) be certain of what your evidence is, (b) be highly confident that it supports ~P, and (c) have high confidence in P anyway. So you are akratic because you are uncertain (or have a false belief) about what E, your evidence, supports. In Dartboard, however, you are akratic because (in part) you are uncertain about what your evidence is, compatible with your being highly confident, or even certain, of the truth about the evidential support relations. So while the two akratic states look similar, they come about through different kinds of uncertainty. 37

The second difference between the cases – and the one I will focus on at greater length – in how we should expect the evidence to vary with the truth of the propositions it supports: Lucy’s guilt in Sleepy Detective, and Ring in Dartboard. 38 In cases like Sleepy Detective, our evidence is usually “truth-guiding” with respect to propositions about the identity of the guilty suspect (and most other propositions, too). By this I mean simply that the evidence usually points to the truth: when it justifies high confidence in a proposition, that proposition is usually true, and when it justifies low confidence in a proposition, that proposition is usually false. If a detective’s first-order evidence points to a particular suspect, that suspect is usually guilty. If it points away from a particular suspect, that suspect is usually innocent.

This feature of Sleepy Detective is by no means exceptional. In most cases, we should expect our evidence to be truth-guiding: otherwise, why should we ever believe what it supports? Moreover, this feature of cases like Sleepy Detective is something that should be available to an agent like Sam: Sam should know, even before looking at his

37 See Titelbaum [ms] for further discussion of this distinction.
38 See Elga [forthcoming] for a more condensed version of this point, which Elga brings up in defense of his New Rational Reflection principle.
evidence, that evidence of this type usually supports the truth about which suspect is guilty.

With this in mind, let’s think about what happens if Sam becomes epistemically akratic in Sleepy Detective, as Level-Splitting suggests. If he is akratic, he cannot continue to believe that his evidence is truth-guiding. Although he will be highly confident that the evidence supports Lucy’s innocence – a condition that would normally go along with her actually being innocent – he will continue to be highly confident of her guilt. Epistemic akrasia seems so odd in this case, in part, because it involves ignoring the reasonable background expectation that one’s evidence is truth-guiding.

In Dartboard, however, the evidence is not truth-guiding, at least with respect to propositions like Ring. Instead, it is falsity-guiding. It supports high confidence in Ring when Ring is false – that is, when the dart landed at <3,3>. And it supports low confidence in Ring when Ring is true – that is, when the dart landed at <3,2>, <2,3>, <4,3>, or <3,4>. This is an unusual feature of Dartboard. And it is only because of this unusual feature that epistemic akrasia seems rational in Dartboard. You should think that you should have low confidence in Ring precisely because you should think Ring is probably true – and because your evidence is falsity-guiding with respect to Ring. Epistemic akrasia is rational precisely because we should take into account background expectations about whether the evidence is likely to be truth-guiding or falsity-guiding.

The reasoning that we just went through has a strange result in Dartboard: it justifies epistemic akrasia. But if we apply that same reasoning to a more standard case like Sleepy Detective, where our evidence is truth-guiding, we get the standard Conciliatory verdict. Returning to that case, suppose that Sam starts off rationally expecting his first-order evidence to be truth-guiding. Suppose he also has higher-order evidence (Alex’s testimony) suggesting that his likelihood of interpreting the first-order evidence correctly is no better than chance. Even after seeing the first-order evidence, then, Sam should remain highly uncertain about what it supports. And because he still expects his evidence to be truth-guiding, his uncertainty about what the evidence supports should bring with it uncertainty about which suspect is guilty. So even after seeing the evidence, Sam should be highly uncertain about which suspect is guilty. For any
particular suspect, including Lucy, he should end up with low credence that that suspect is guilty. And this is exactly what the Conciliatory view recommends.

Now let’s return to some of the problems for epistemic akrasia, and see how these arguments work in Dartboard. First, let’s consider Sam’s conclusion that his evidence is misleading in Sleepy Detective (from section 3, above). I argued above that it is irrational for Sam to believe that his evidence is misleading, in part, because it seems that he can identify something he could do to avoid being misled. In other cases, however, believing that one has misleading evidence might not be so bad. Specifically, it did not seem at all problematic in cases like the lottery, where there is nothing one can do to avoid being misled. In Dartboard, it does seem that you should think that your evidence is misleading. You should, after all, have .8 credence in Ring, but you should also be highly confident that your credence in Ring should be .2. But if we look more closely at the details of Dartboard, we can see that it might be best treated as a case like the lottery, where one cannot avoid being misled.

Here is why: in Dartboard, you should be .8 confident that the dart landed on one of <3,2>, <2,3>, <4,3>, or <3,4>, but you have no idea which. Each of those evidential situations rationalizes a different credence distribution. And although those different credence distributions each assign .2 credence to Ring, they disagree about nearly everything else. For example, if the dart landed on <3,2>, you should have .2 credence in Ring because you should have .2 credence that the dart landed on <3,2>. But if the dart landed on <3,4>, you should have .2 credence in Ring because you should have .2 credence in <3,4> – and 0 credence in <3,2>! So although you should be .8 confident that you should have .2 confidence in Ring, there is no particular credence distribution that you think you should adopt. While you should think your evidence is misleading, this belief nevertheless seems stable.

Second, let’s return to the bootstrapping argument. I argued that if Sam becomes epistemically akratic in a long string of cases like Sleepy Detective, he will find himself in an unusual epistemic situation – one that requires explanation. The best explanation for that situation is that he was never seriously rationally impaired in the first place. So, I argued, Sam should bootstrap.
The bootstrapping argument relies on the fact that in cases like Sleepy Detective, we rationally expect our evidence to be truth-guiding. This is why believing irrationally is usually unreliable, why finding oneself epistemically akratic is unusual, and why believing rationally is a good explanation for one’s having true beliefs. But when one’s evidence is falsity-guiding, as it is in Dartboard, things may be different. Finding oneself with true-but-irrational beliefs is to be expected in cases like Dartboard. And because one’s evidence is not truth-guiding, believing rationally is not a good explanation for having true beliefs about propositions like Ring. In Dartboard, then, the bootstrapping argument does not get off the ground.

If this type of Conciliatory view is right, we cannot endorse the Non-Akrasia Constraint as a universal prohibition against a certain combination of attitudes. That’s because it might sometimes be rational to have high confidence in both “P” and “my evidence doesn’t support P”. But contra Level-Splitting, this combination of attitudes is not made rational merely by our receiving misleading higher-order evidence. Instead, epistemic akrasia will only be rational in those odd cases where we expect our evidence to be falsity-guiding. The next task for this type of Conciliatory view, then, is to develop a more nuanced account of the interaction between epistemic levels: one that takes into account whether we rationally expect our evidence to point to the truth. 39

Conclusion
I have argued that Level-Splitting views – those on which epistemic akrasia can be rational in cases like Sleepy Detective – incur substantial intuitive costs. This is good evidence for the Non-Akrasia Constraint: the problems for Level-Splitting suggest that at least in paradigmatic cases, it cannot be rational to have high confidence in both “P” and “my evidence doesn’t support P”.

Some important questions remain. First, it remains to be seen how higher-order defeat might work: what if, as Lasonen-Aarnio argues, there is no satisfactory, coherent way of spelling this out? And second, much more should be said about when there is

39 Elga's “New Rational Reflection” principle is one promising attempt to do just that. So, there is reason to be optimistic about the prospects for this type of Conciliatory view. See Elga [forthcoming].
higher-order defeat: when should higher-order evidence affect our first-order beliefs? And when, if ever, can epistemic akrasia be rational?

Though the discussion here doesn’t settle these questions once and for all, we do now know something about how a good theory of higher-order evidence will answer them. First, in order to be extensionally adequate, a good theory of higher-order evidence should say that, in paradigmatic cases like Sleepy Detective, epistemic akrasia cannot be rational. That’s because in those cases, epistemic akrasia leads to bad reasoning and irrational action. And we should take these problems seriously even in the absence of general principles: it is difficult to imagine even a rough picture of the relationship between epistemic levels that would vindicate a pro-akrasia verdict in these paradigmatic cases without committing us to absurd consequences.

Second, we now know something about why higher-order evidence should affect our first-order attitudes in paradigmatic cases of misleading higher-order evidence. That is: in normal cases, we rationally expect our evidence to be truth-guiding. And when we expect our evidence to be truth-guiding, what we should believe about the world should line up with what we believe our evidence supports. We should not have high confidence in both “P” and “my evidence doesn’t support P”.


Your little brother’s birthday is coming up. You want to give him a present, but you’re not sure what would be best: it’s got to be something he doesn’t already have, and it’s got to be good. As you’re walking along one day, contemplating your options, a mysterious stranger approaches and hands you a pamphlet. It reads: *Give the gift of epistemic rationality! Our experienced technicians will gently scramble your friend or loved one’s brains so that, from now on, she will believe what her evidence supports. Results guaranteed. All previously acquired evidence will be maintained. Call today!* “There’s my answer!” you think to yourself. After all, you know that your brother – unlike you – often falls short of the epistemic ideals. By making him epistemically rational, this spa treatment should thereby make him epistemically better off.

Believing rationally is epistemically valuable, or so we tend to think. It’s something we strive for in our own beliefs, and we criticize others for falling short of it. We theorize about rationality, in part, because we want to be rational. But why? I’ll call this the “value question”. Before picking up the phone, it seems, you should have an answer to the value question: you should be able to explain how believing rationally would benefit your brother. The answer to this question should be provided by the right epistemological theory. A good epistemological theory should allow us to identify a special feature of rational belief – something that your irrational brother now lacks, but would gain only by becoming rational. And that feature should be something worth caring about as much as we care about believing rationally.

The way one answers the value question will depend on the details of one’s view. This paper explores one important way in which those details matter: whether, and to what extent, rationality is *permissive*. (That is, does our evidence completely determine what we should believe? Or do we have some leeway?) In the first half of the paper, I’ll set out and compare the ways in which *extreme permissivism* and *impermissivism* can answer the value question. In the second half, I’ll turn to *moderate permissivism*. The surprising upshot of this discussion is that moderate permissivism, despite its popularity and prima facie appeal, fares particularly badly when evaluated from this perspective.
Considering the value question therefore gives us a new reason to worry about moderate permisivism.

1. Extreme Permissivism

Before getting into the details of any particular view, I should say a bit more about what I mean by “permissivism” and “impermissivism”. Impermissivism is the view that, given a total body of evidence, there is a unique ideally rational doxastic response that one can take to any proposition.\(^{40}\) Permissivism is simply the denial of impermissivism, and encompasses a wide variety of diverse views. Some permissivists – I’ll call them “extreme” permissivists – hold that any way of responding to evidence is rationally permissible, as long as one maintains certain formal relationships among one’s beliefs. Others – “moderate” permissivists – hold that only some ways of accommodating evidence are rationally permissible. What these views have in common is the thought that two agents can share all of their evidence, but nevertheless have different rational beliefs.

Questions about permissivism largely cross-cut those about internalism and externalism, synchronic versus diachronic notions of rationality, and many others. One could hold a variety of different positions regarding these questions and still face a further choice between permissivism and impermissivism. Since the arguments here won’t hinge on these other issues, it will help to restrict our attention to views that agree on everything aside from their stance on permissivism. To that end, I’ll assume for present purposes that we have settled on a common conception of what counts as “evidence” (understood as what rational agents “learn”), and a common way of spelling out formal requirements of coherence or consistency.

The Bayesian framework provides a good way to compare permissive and impermissive views while holding everything else constant.\(^{41}\) Bayesian views agree on certain formal constraints: rational agents must have probabilistically coherent credences, and update on new evidence by conditionalization. An agent’s rational belief state at any time depends on both her total evidence and on her initial credence function (her prior

\(^{40}\) This is often called “Uniqueness”. See White [2007], Feldman [2007], Christensen [2007], Kelly [forthcoming], Cohen [forthcoming], and Schoenfield [forthcoming] for various formulations of the thesis.

\(^{41}\) Bayesianism will also provide a natural setting for the discussion here; many people find my main target view, moderate permissivism, much more compelling for credences than for full belief.
unconditional and conditional credences, or “priors”). Different priors encode different ways of reasoning inductively; so, two agents who share their evidence, but not their priors, might disagree about quite a bit. Bayesianism by itself is silent on the question of which priors are rationally permissible, and is therefore compatible with a wide range of positions on permisivism. As I’ll carve up the territory, extreme permisivists hold that any coherent priors are rationally permissible; moderate permisivists hold that several, though not all, coherent priors are rationally permissible; impermisivists hold that there is just one set of rationally permissible priors.42

With the debate set out this way, we can now look at how extreme permisivism – in particular, (extreme) Subjective Bayesianism – answers the value question. According to Subjective Bayesianism, we are rationally required to follow the Bayesian formal constraints (probabilistic coherence and conditionalization). But there are no further requirements dictating what our priors should look like; any are rationally permissible.

If Subjective Bayesianism is right, why does rationality matter? Recent work in formal epistemology has brought out one way in which Subjective Bayesians can answer this question. That is, they can point out that rationality, on their view, is both necessary and sufficient for complying with the following principle:

**Immodesty:** The credences recommended by your own epistemic rule, given a body of evidence, should uniquely maximize expected accuracy for you.

In a Bayesian context, following your “epistemic rule” amounts to updating your prior credence function by conditionalization; Immodesty dictates the way you should regard that epistemic rule, if you compare it to others. I’ll say a bit about what Immodesty means, and then come back to its relation to Subjective Bayesianism.

The main motivation for Immodesty is the thought that a rational agent should be doing well by her own lights, in a particular way: roughly speaking, she should follow the epistemic rule that she rationally takes to be most truth-conducive.43 It would be irrational, the thought goes, to regard some epistemic rule as more truth-conducive than

---

42 My uses of “extreme” and “moderate” in this context follow Meacham [ms]. White [2007] uses the same terms to mark a different distinction.

43 I don’t mean to build in much by “rule” here; we can think of an epistemic rule as a mapping from evidence to belief states. “Following” a rule, for present purposes, just involves adopting the belief state that corresponds to your total evidence.
one’s own, but not adopt it. (Of course, we have to be careful with scope here: it’s often rational to believe that there exists some epistemic rule, or way of responding to evidence, that is more truth-conducive than one’s own. A rule that ignored misleading evidence, for example, would be more truth-conducive than a rule that took all evidence into account; it seems perfectly rational to believe that there are such rules out there, though we can’t identify which ones they are. What’s not rational, then, is to regard some particular epistemic rule as more truth-conducive than one’s own, while also knowing which belief state that rule recommends in every case.) Immodesty says that, among particular epistemic rules that you can compare to your own, you should regard your own as optimal, or as giving you the best shot at having true beliefs. An immodest agent’s beliefs are therefore stable in a certain way: we might say that she’s living up to her own epistemic standards, or taking (what she sees as) the best epistemic means to her end of believing truly. This seems like a good thing.

I put the motivation for Immodesty in terms of truth-conduciveness, but of course credences can’t be true or false. So when we’re working with credences, we can think instead about accuracy, or closeness to the truth. If P is true, a credence of .8 in P is intuitively more accurate than a credence of .7; the opposite holds if P is false. Formal epistemologists often measure accuracy using utility functions (called “scoring rules”) that assign value to credences depending on how close they are to the truth. The more accurate your credence, the better the score will be. When you aren’t sure how accurate some credence in P is – which will happen whenever you aren’t sure whether P is true – you can find the expected accuracy of that credence by taking a weighted average of the credence’s score in cases where P is true and in cases where P is false; weights are given by your own credence in P and in ~P, respectively. In order to be immodest when you aren’t sure what’s true, you must regard the credences recommended by your own epistemic rule as maximizing expected accuracy.44

44 To head off a couple of worries here: you might be concerned that, because it dictates how one should regard one’s own beliefs and epistemic methods, Immodesty implausibly requires us to form beliefs about all of these matters. To avoid that, we can think of Immodesty as a principle of propositional, rather than doxastic justification: it tells us what opinions would make sense for a rational agent to have, should she happen to form opinions at all. Second, one might worry that, since it’s framed in terms of maximizing a certain value, Immodesty appears to commit us to consequentialism about epistemic rationality. For reasons of space, I won’t get into this issue in depth here. But it’s not clear to me that non-consequentialists should be concerned that the Immodesty demand requires consequentialism. In general, we can think about
Back to Subjective Bayesianism: on that view, why is rationality necessary and sufficient for obeying Immodesty? Formal epistemologists have defended each of the Bayesian formal constraints, separately, by appealing to Immodesty (or closely related principles). Joyce argues that if you have *coherent* credences, you will be immodest: from your perspective, your own credences will maximize expected accuracy. But if you have *incoherent* credences, you will not be immodest. Greaves and Wallace defend Subjective Bayesianism's second formal requirement, conditionalization, by appealing to similar considerations of expected accuracy. If you have coherent credences, they argue, then from the point of view of your own credences, updating by conditionalization will maximize expected accuracy (and other methods of updating will not). Putting those together, a Subjective Bayesian can argue that a rational believer is just an immodest one.

Immodesty is widely held to be a necessary condition on rational belief. And if Joyce’s and Greaves and Wallace’s formal arguments are right, obeying the Bayesian
formal constraints will guarantee that one complies with Immodesty. This means that Immodesty will be true on any Bayesian view, including those that are less permissive. But for Subjective Bayesians, who only endorse the formal requirements, believing rationally just is believing immodestly. Rationality is good because it’s good to do well from one’s own perspective; you should make your little brother rational because it will at least guarantee that he will follow an epistemic rule that does well by its own lights.

2. Impermissivism

Insofar as Immodesty is a good thing, (extreme) Subjective Bayesianism seems to provide a good answer to our central question: it points to a unique feature of rational belief that is plausibly worth caring about. But many people worry that the notion of rationality articulated by Subjective Bayesianism, and by other extremely permissive views, is simply too weak. The most obvious (and notorious) reason for doubt is that if we only place formal constraints on rational belief, we have to count all kinds of internally consistent views (skeptics, grue-projectors, Tea Party “birthers”, etc.) as rational. One might argue that this is simply implausible: holding those views, at least given the kind of evidence that most of us have, is not rational. So Subjective Bayesianism just doesn’t give us a good account of what rationality is like.

Second, one might take issue with the Subjective Bayesian’s answer to the value question: though Immodesty is a good thing, perhaps it’s not good enough to explain why we should care about rationality. (In particular, one might add, it’s not obvious how good it is to be immodest if, as Subjective Bayesianism allows, one’s rational beliefs could so easily be so radically mistaken.)

Epistemologists who take these concerns seriously argue that in addition to formal constraints, there are also strong substantive constraints on rational belief. Impermissive views hold that the substantive constraints are so strong that rational permission and rational obligation coincide. According to (Extreme) Objective Bayesianism, which we can take as a paradigmatic impermissive view, there is just one rational set of priors. A rational agent must have those priors, and update on new evidence by conditionalization.

believing optimally, in some sense – but objects to interpreting it in terms of expected accuracy. Many epistemologists working in the Bayesian framework take up Immodesty, spelled out in expected accuracy terms, as a datum; one clear example is Moss [2011].
So two agents with the same total evidence should, rationally, have the same total belief state.\textsuperscript{48}

Impermissivism is attractive for a number of reasons. Unlike extreme permissivism, it allows us to rule out skeptics, counterinductivists, and grue-projectors, and defend other substantive rational requirements traditionally discussed by epistemologists. It is also broadly in line with how we often think about evidence: we talk about “what the evidence supports” as if there is only one evidential support relation, and we ask what a rational agent would believe under certain circumstances as if there is only one option for what that could be.

Impermissivism also offers an attractive answer to the value question. That is, impermissivists can argue that rationality matters because it guides us to the truth, in a particular way: given a body of evidence, the rational credences are those that maximize expected accuracy. To see why, suppose you know that your way of accommodating evidence is rational, and you know that rationality is impermissive. Yours is therefore the only rational way of responding to evidence, so you know that any rational believer will accommodate her evidence by conditionalizing on your priors. And because of Immodesty – which holds on any Bayesian view – those credences should maximize expected accuracy from the point of view of your credences. Putting these pieces together, you can explain why rationality is valuable: believing as rationality recommends maximizes expected accuracy.\textsuperscript{49}

To put the point another way, if impermissivism is true, a rational agent who knows she’s rational, and knows what rationality requires, should be able to make the following argument:

\begin{center}
Where \( E \) is any body of total evidence, and \( C \) is any credence function:
\end{center}

\textsuperscript{48} See White [2007] for a defense of impermissivism. See also Feldman [2005] and Christensen [2007]. Williamson [2000]’s notion of “evidential probability” is a version of Objective Bayesianism, though its relation to justified beliefs or degrees of confidence is not straightforward.\textsuperscript{49} Dogramaci [2012] makes a similar point, in explaining why we should defer to testimony from other rational agents: “...[Y]our own beliefs can serve as bases for inferred conclusions that I can then acquire by testimony. And this is all possible because, when we share rules, I can trust that you will draw the same conclusion from an evidential basis that I would.” (p. 524) Dogramaci does not explicitly endorse either impermissivism or permissivism, but his defense of “epistemic communism” is similar to the account of epistemic value that I offer on behalf of the impermissivist.
P1. If C is any rationally permissible response to E, then my epistemic rule will recommend C, given E.
P2. If my epistemic rule recommends C, given E, then C maximizes expected accuracy given E.
C. If C is a rationally permissible response to E, then C maximizes expected accuracy given E.

A rational agent should hold P1 just by virtue of knowing that she’s rational, and knowing that rationality is impermissive. She should hold P2 because of Immodesty. From those two, she can conclude that any rational response to total evidence E will maximize expected accuracy. The benefits of believing rationally are exactly the same benefits that one receives from responding to evidence as she does. Moreover, drawing this conclusion is no accident. By virtue of our taking up the perspective we did – that of a rational agent who knows what she believes, and knows what rationality requires – we are guaranteed to reach this conclusion.

Subjective Bayesianism merely requires that believers do well from their own perspectives (whatever that might involve); if you make your little brother rational by Subjective Bayesian standards, you will only be able to guarantee that he maximize expected accuracy by his own lights. His expected accuracy by your lights could be way off, even if you share all of your evidence. So, for Subjective Bayesians, rational agents should endorse their own credences, and their own epistemic rules, but shouldn’t endorse other rational agents’ epistemic rules. Objective Bayesianism gives us something much stronger: if you make your brother rational by Objective Bayesian standards, you can explain that this is a good idea because it guarantees that he will maximize expected accuracy, full stop. Rational Objective Bayesians should endorse their own epistemic rules and those of others in just the same sense.50

50 While I’m calling this a “truth-guiding” account of the value of rationality, there’s a sense in which it isn’t “really” truth-guiding; the connection to truth is cast in subjective, rather than objective, terms. So it’s possible on this account to have a rational agent who is vastly mistaken about the world. An impermissivist should not say that rationality guarantees that this is not our situation. But she can say that such a situation is very unlikely. Cohen [1984] raises some worries for a subjective connection between rationality and truth, mainly targeting the view that one must believe that one’s epistemic rules are reliable in order to use them. Requiring us to form all of these higher-order beliefs, Cohen argues, is an unrealistic and undue cognitive burden. We can sidestep many of these worries by thinking of this view as one about propositional, rather than doxastic justification: the idea here is that if one were to form beliefs about the reliability of one’s methods, one would be justified in taking the attitudes described by Immodesty.
3. Moderate permissivism

Impermissivism gave a good answer to the value question: if impermissivism is right, it's clear that rationality is worth caring about. But some epistemologists worry that the view is too demanding. If there is just one rational way of accommodating evidence, there is no possibility of rational disagreement or leeway in what we may believe. Some argue that disagreement among scientists or jurors with shared evidence does not mean that one party or another must be making an epistemic error. It might also seem implausible to think that rationality extends to all subject matters, or that rationality narrows down the permissible responses to a body of evidence as drastically as impermissivists suggest. Kelly describes the current state of play as follows:

Notably, even Bayesians who are considered Hard Liners for holding that there are substantive constraints on rational prior probability distributions other than mere probabilistic coherence typically want nothing to do with the suggestion there is some uniquely rational distribution. With respect to this long running debate then, commitment to [impermissivism] yields a view that would be considered by many to be beyond the pale, too Hard Line even for the taste of most Hard Liners themselves.

While extreme permisivism seemed too weak, impermissivism seems too strong. In light of this, many epistemologists are drawn to an intermediate position:

moderate permissivism. Recently, some epistemologists have explicitly defended this type of view. And it also seems to enjoy something of a “default” status among others. In many ways, moderate permissivism looks like a happy compromise between extreme permissivism and impermissivism. Like impermissivism, moderate permissivism allows us to place substantive constraints on rational belief. (For example, a moderate permissivist might insist that rational agents assign comparatively higher initial credence to non-skeptical hypotheses, and comparatively lower initial credence to skeptical ones.)

---

51 See, e.g., Rosen [2000], Douven [2009], Cohen [forthcoming], Kelly [forthcoming], Teller [forthcoming], Schoenfield [forthcoming], and Meacham [ms] for objections to impermissivism along these lines.

52 Kelly [2010], p. 11. Douven makes a similar observation: “[M]ost Bayesians nowadays think rational degrees of belief are to satisfy additional constraints [beyond satisfying the axioms of probability]... Still, to the best of my knowledge no one calling him- or herself a Bayesian thinks that we could reasonably impose additional constraints that would fix a unique degrees-of-belief function to be adopted by any rational person and would thereby turn Bayesianism into a fully objective confirmation theory.” (Douven [2009], p. 348)

53 See especially Kelly [forthcoming] and Schoenfield [forthcoming] for examples of the types of view I have in mind.
But like extreme permissivism, moderate permissivism also allows us to make sense of rational disagreement, and gives us some leeway in how we may respond to our evidence.\textsuperscript{54}

In explaining why you should want your little brother to be rational, moderate permissivists cannot help themselves to the accounts offered by the other two views. They cannot appeal only to the value of complying with Immodesty, because for moderate permissivism, rationality requires more than just coherence and conditionalization. (And anyway, moderate permissivism is designed to \textit{rule out} many immodest believers.) The impermissivist’s strong accuracy-based answer is unavailable as well: making your brother rational will not guarantee that he end up with your priors, so if he becomes rational, he will not be guaranteed to maximize expected accuracy from your point of view. (In the argument in the last section, moderate permissivists must deny P1.)\textsuperscript{55}

But even with impermissivists’ answer out of reach, it’s plausible that moderate permissivists’ answer to the value question should involve some connection to truth or accuracy. One of the motivations for rejecting \textit{extreme} permissivism, after all, was the thought that radically out-of-touch believers should not count as rational. So a weaker accuracy-based strategy looks like a promising option for moderate permissivists: although rationality does not \textit{maximize} expected accuracy, perhaps it \textit{increases} expected accuracy. Rational believers should, in general, do better accuracy-wise than irrational

\textsuperscript{54} Moderately permissive views can be more or less moderate. Pettigrew [2012] and Meacham [ms], for example, discuss versions of moderate permissivism on which the only substantive constraint is compliance with the Principal Principle. This proposal is very close to Subjective Bayesianism, and Pettigrew argues that it can be defended on similar lines. It might turn out, then, that this view can answer the value question along similar lines as well. Because it is so permissive, though, it will not be attractive to anyone who is worried about things like skepticism and grue-projection. For the present discussion, I will set aside this type of view and focus on less-permissive moderate views.

\textsuperscript{55} It’s true that moderate permissivists can’t say that rational belief maximizes expected accuracy as assessed from \textit{the} rational perspective. But couldn’t they say, instead, that rational belief maximizes expected accuracy from \textit{a} rational perspective? (Each rational credence function will maximize expected accuracy relative to one rational credence function: itself.) Moderate permissivists could say this, but they also will have to say more: \textit{extreme} permissivists, after all, can also say that rational belief maximizes expected accuracy as assessed from a rational perspective. So moderate permissivists will need to say why their view goes beyond extreme permissivism. Once we add something to this view, however, our answer to the value question becomes less unified. If we wanted to identify a unique, valuable property of rational belief, this kind of strategy won’t give us one. Though that isn’t a knock-down objection, I think it is at least reason to worry that this strategy won’t yield a satisfactory answer to the value question. Thanks to Dennis Whitcomb for suggesting this response on behalf of moderate permissivism, and for helpful discussion on this point.
believers. This strategy, if successful, would be a good start to answering the value question. Making your brother rational would make him epistemically better off because it would give him a better shot at believing accurately.

This section will be primarily devoted to examining the prospects for giving an “increased expected accuracy” answer to the value question. I’ll argue that the strategy faces serious challenges; looking at why will illustrate some more general ways in which moderate permissivism yields odd results when we consider how rational agents should think about rationality itself. At the end of the section, I will return to the question of whether other kinds of answers to the value question might work.

The increased expected accuracy answer is initially attractive. A quick-and-dirty argument supporting the strategy might go like this: according to moderate permissivism, there are some bodies of evidence (let’s call one “E”) such that, although rationality does not mandate a unique doxastic response, it does require that one’s credence in P fall within a certain range. Now consider how a rational agent with total evidence E, whose credence in P falls within the rational range, should regard others with total evidence E who do and do not fall within that range. Because of Immodesty, she should take her own credence in P to maximize expected accuracy; the expected accuracy of other credences in P will get higher and higher, from her point of view, as they get closer to hers. Other credences in P that fall within the range will generally be closer to hers than credences outside the range. So other agents’ credences will tend to have higher expected accuracy if they are rational, and lower expected accuracy if they are not.

But while the increased expected accuracy strategy looks good at first glance, closer inspection reveals some problems, which I’ll outline below: First, this view runs into problems with justifying coherence. Second, even if we set that aside, the strategy does not succeed in picking out a unique positive feature of rational belief. Following this

56 Though this position isn’t explicitly endorsed by permissivists as far as I can tell, it is often in the background. For example, Kelly writes, “the Permissivist might think that what permissive cases there are, aren’t all that permissive. ... [Suppose] you and I agree on the basis of our common evidence that the Democrat is more likely than not to be elected. ... The only difference between us is this: [you] give a bit less credence to the proposition that the Democrat will win than I do. Here there seems little pressure for me to conclude that you are less reasonable than I am.” (Kelly [forthcoming], p. 2-3)

57 This is an intuitively plausible constraint on acceptable ways of measuring both closeness and accuracy of credence functions. For a proof that this constraint is true of strictly proper scoring rules, see Pettigrew [2013] (appendix).
strategy also allows rational agents to disagree about the answer to the value question, which I’ll argue is an odd and undesirable consequence. Finally, I’ll argue that any expected-accuracy-based account puts moderate permissivism at a disadvantage in comparison to impermissivism.

To see how the first problem arises, consider what a strong version of the increased expected accuracy strategy would say: that rationality guarantees a certain high degree of expected accuracy, from the point of view of any rational agent. That is, that given a body of evidence, there is an “expected accuracy threshold” that divides the rational belief states from the irrational ones. If your brother is irrational, he’s below the threshold; making him rational would ensure that he end up above the threshold. We might think of this as a “satisficing” view, in contrast to the impermissivist’s maximizing view.

Now suppose you’re rational, and you have probabilistically coherent credences. You want your brother to meet the expected accuracy threshold, so you need to give him credences closer to yours. But do you need to give him coherent credences? Not if you only care about expected accuracy: in addition to the coherent credences that meet the threshold, there will be lots of incoherent credences that do too. So if moderate permissivists want to require coherence as well as high expected accuracy, they will need to say something more: the satisficing view can’t give us the whole story.

Even setting aside those worries, however, the satisficing view is still false. Suppose a body of evidence, E, rationalizes any credence in P from .6 to .8. Alice, Bob, and Charlie all have total evidence E, and different credences in P, as follows:

<table>
<thead>
<tr>
<th>Credence in P, given E:</th>
<th>[ rational range ]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.6 .7 .8 .81</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>C</td>
<td>---------------------</td>
</tr>
</tbody>
</table>

58 For example, suppose your credence in P is .5, and your credence in ~P is also .5. It might be rational to have .6 credence in P, and .4 in ~P — suppose that coherent assignment of credences has high expected accuracy as assessed by your credences and your scoring rule, so it meets the required threshold. But then some incoherent assignments of credences will plausibly also meet the threshold: for instance, .51 credence in P and .49 credence in ~P. Thanks to Jennifer Carr and Brian Hedden.
Suppose further (as we have assumed so far) that our rational agents know everything we do about their setup: Alice and Bob both know what they believe, and know what rationality requires. How should Alice and Bob answer the value question?

From Alice’s perspective, something like the satisficing view makes sense: all of the rational credences in P are closer to her own than the irrational credences in P, so she should regard the rational credences as having higher expected accuracy than the irrational ones. Speaking loosely, Alice’s beliefs about what’s rational line up with her beliefs about what’s accurate. But from Bob’s point of view, things look much worse. Imagine how Bob should compare Alice and Charlie’s beliefs. “Alice is rational,” he might think to himself; “her credence is in the permitted range. Of course, it’s probably a bit less accurate than mine; so, I’m doing somewhat better than she is in terms of accuracy. But nevertheless, since she’s rational, there’s something commendable about her having credence .7 in P.” Then he considers Charlie: “Charlie is irrational; his credence is outside the permitted range. So he is missing out on that special commendability that Alice has. But fortunately for him, he’s doing quite well, accuracy-wise. In fact, he seems to be doing better than Alice!”

While many permissivists accept that there can be cases like this, some do not. For instance, Cohen [ms] argues that rationality is only permissive when one is unaware of the other permissible options on one’s evidence. A view like this might be able to avoid many of the problems I raise here for moderate permissivism. I don’t have the space to give this view the attention it deserves, but I will mention a few reasons to think that permissivists should be hesitant to adopt it. First of all, Cohen’s view commits us to the claim that there can be widespread rational ignorance, or rational false beliefs, about what rationality requires: and indeed, that this kind of ignorance is rationally required in all permissive cases. This is a very strong conclusion (as Ballantyne and Coffman [2012] point out). Second, this view undermines some popular motivations for permissivism: for example, it implies that you and I could never rationally recognize our own situation as a “reasonable disagreement” of the type supposedly found on juries or among scientists.

Cohen embraces this conclusion, and argues that when we find out which other credences are rational (perhaps through disagreement with others) we should conciliate. But this means that if a rational agent takes any view of her own epistemic situation at all, she must believe that if her evidence is permissive, her own credence in response to it is near the middle of the permissible range (at least in expectation). (If she thought her credence was near the lower bound of the permissible range, say, she would have pressure to move towards the middle.) Why is the middle of the range so special, and why should we try to approach it in response to disagreement? See Cohen [forthcoming] and Christensen [2009] for further comments. Thanks to Louis deRosset for helpful discussion on this topic.

I’m simplifying a little here; depending on which scoring rule we’re using here, it might turn out that the “middle” of the rational range isn’t actually where we should put Alice in order to make the threshold view come out true from her point of view.

Bob should think that his own beliefs maximize expected accuracy, again, because of Immodesty. Schoenfield [forthcoming] argues that this is why someone in Bob’s situation should not regard the choice between his credence and Alice’s as “arbitrary”, contra White [2007]; we should stick to our own credences in permissive cases because we see those credences as maximizing expected accuracy.
The satisficing view predicts that rational agents should regard all rational responses to E as having higher expected accuracy than all irrational ones. But in this case, that wasn’t true: since Charlie’s credence in P is much closer to Bob’s than Alice’s is, Charlie’s credence in P has higher expected accuracy, from Bob’s point of view, than Alice’s. This observation makes the value question especially salient. From Bob’s point of view – and in fact, from the point of view of any rational agent, other than Alice – what is the sense in which rational beliefs are better than irrational ones? What is it that rationality gives Alice, but that Charlie lacks? So far it’s still not clear.

Maybe there’s a way to weaken the increased-expected-accuracy view in response to this first objection. For instance, we could say that rationality increases expected accuracy “in general”, or “on average”; our next job would be to say more about what that means. But even if we can make sense of a weaker view, the increased expected accuracy account still faces additional challenges.

The third problem with the increased expected accuracy strategy is that it allows different agents to rationally disagree about the value of rationality. Staying with the example above, compare how Alice and Bob will each assess the extent to which rationality tends to increase one’s expected accuracy. For Alice, whose credence in P is in the middle of the permissible range, believing rationally is a reliable way of believing something close to what she believes, and therefore a reliable way of gaining high expected accuracy. But for Bob, whose credence in P is on the outer edge of the permissible range, believing rationally will not appear to be a particularly good way of gaining high expected accuracy. So Alice and Bob can rationally disagree about how valuable rationality is.

Of course, for permissivists, rational disagreement happens all the time. But this instance of rational disagreement seems especially odd. For one, what would settle Alice and Bob’s disagreement? If we say that one of them is right – for example, if we argue that Alice’s position is authoritative, and Bob is mistaken – we seem to undermine one of the main reasons to hold permissivism in the first place. Permissivism is partly motivated by precisely the thought that there is no unique, privileged way to respond to one’s

---

62 Again, I’m glossing over some issues about “closeness” here (see fn. 71). But I think we can make sense of the general point without going into detail on this issue.
evidence. Designating one particular response to evidence as the unique arbiter of an (arguably) a priori matter seems to give up that thought. Taking this line would also require us to say that poor Bob is doomed to have false (but justified) beliefs about the value question; this might also seem like an odd consequence if we take the value question to be a priori.

Alternatively, we could claim that there is just no fact of the matter about the degree to which rationality increases expected accuracy, and hence, no unique right answer to the value question. Alice and Bob’s answers are equally good. (So, we would give a kind of relativist account of rational value.) But if we go this way, there’s still something puzzling about comparing the two answers: while Alice’s perspective on rationality seems relatively sensible and coherent, Bob’s does not. From Alice’s perspective, moderate permissivism might look quite attractive. Field, for instance, argues (briefly) that we should be “moderate relativists” about epistemic rationality because we should recognize several different ways of responding to evidence, even those that differ from our own, as “by no means beyond the pale”. For Alice, this makes sense: Alice should regard herself as a reasonable, middle-of-the-road believer, and should regard other rational agents as similarly reasonable.

But just as we saw with the satisficing view, Field’s proposal doesn’t make sense for someone in Bob’s position. If Bob agrees with Alice about which epistemic rules are “beyond the pale”, he should see himself as being just on the edge of craziness; this seems like an unstable position to be in. But if Bob disagrees with Alice about which epistemic rules are beyond the pale, it seems like Bob should also disagree with Alice about which epistemic rules are rational. (This leads to more trouble, along the lines of what we saw above. If Alice and Bob disagree about what’s rational, what settles that disagreement? And are some agents, like Bob, doomed to have rational false beliefs about what’s rationally required?) However we spell it out, it looks like moderate permissivists will have trouble explaining Bob’s predicament.

Finally, even supposing that we can answer these objections, there is something unsatisfying about any increased expected accuracy answer to the value question. That is: it looks like anything moderate permissivism can do, expected-accuracy-wise,

---

63 Field [2000], p. 141.
impermissivism can do better. Suppose you have two options for your brother's birthday present: spa #1 makes him rational, and spa #2 gives him *your particular* rational epistemic rules. If you care about your brother's accuracy, it might be better to send him to spa #1 than to leave his irrational epistemic rules in place. But it would an *even* better idea, for the same reasons, to send him to spa #2. It's hard to see why the first spa is worthy of special recognition at all.

How much trouble will moderate permissivism be in, if it cannot give an accuracy-based answer to the value question? Some might be tempted to respond that this isn't a worry at all: this is just the kind of question that doesn't need to be answered in the first place. But this response won't be satisfying for anyone who *does* want an answer to the value question. And given that extreme permissivists and impermissivists can give accounts of why rationality matters, the claim that the value question is "unanswerable" becomes less plausible.

Assuming that moderate permissivists do want to answer the value question, perhaps they could take some other strategy, completely unrelated to accuracy. Or perhaps moderate permissivists could offer a more complicated account (such as, "rationality is about Immodesty *and* following induction"). But there's reason to think that an accuracy- or truth-based answer is moderate permissivism's best hope.

In answering the value question, we aimed to find some feature of rational belief that explains why rationality is worth caring about. And there *is* something that explains why we should care about each of the individual substantive rational requirements traditionally discussed by epistemologists: induction, anti-skepticism, green-projection, etc. Following these requirements keeps us connected to the world, and helps us get to the truth. (This is a big part of why some epistemologists are so invested in defending substantive requirements in the first place: skeptics and grue-projectors are just getting things wrong!) So a theory of rationality that holds all of these requirements should be able to appeal to their common feature in explaining why these requirements are worth caring about. If moderate permissivism can't appeal to truth, the view misses out on a particularly intuitive, unified account of the value of rationality.
For that reason, it seems unlikely that moderate permissivists will be able to answer the value question in a way that does not appeal to truth at all. The best strategy, for moderate permissivists, might be to fix up some version of the increased expected accuracy answer. But how?

4. A final objection

I've suggested some ways in which extreme permissivists and impermissivists can answer the value question, and given reasons to worry that moderate permissivists may not be able to do the same. But you might worry that the impermissivist’s answer to the value question isn’t so great. In order to give the impermissivist’s answer, we must take up the perspective of someone who already holds a particular impermissive view. (For example, in order to accept an argument that following induction maximizes expected accuracy, it seems like we must already be committed to following induction ourselves.) So, the impermissivist’s answer to the value question is self-supporting.⁶⁴

We can answer this objection, first, by thinking back to how our challenge was originally set up. We started off by observing that we want to be rational, and that if we could, we would want to make others rational too. The challenge was to explain why this is; in looking for an explanation, we wanted an account that a rational agent could give just by reflecting on the nature of the rational requirements according to various different views. But in order to see what an adherent of some view could say in her own defense, we have to see how things look from her perspective. So with the project set up in this way, it’s no surprise that the answers we ended up with turned out to be self-supporting. Self-supporting answers are exactly what we were looking for.

Perhaps the objector wanted something stronger: an account of the value of rationality that appealed to neutral premises, and could convince opponents. But while this more ambitious kind of account would be nice to have, it might turn out that in order to defend any substantial epistemological view, we need to take a much more modest

⁶⁴ Meacham [ms] and Teller [forthcoming] each raise versions of this worry. Similar complaints have been raised against the extreme permissivist’s defense of Immodesty. Subjective Bayesians argue that coherent agents should regard their credences as maximizing expected accuracy according to a proper scoring rule – but why should we think that this is the right way to measure accuracy? One of the main motivations for using proper scoring rules is that they allow coherent agents to be immodest. So the extreme permissivist’s answer to the value question might be similarly unconvincing to those who don’t already hold the view. See Maher [2002] and Gibbard [2008] for two versions of this worry about Subjective Bayesianism.
approach. We might only be able to explain the position from inside, rather than justify it to outsiders. If that's the case, a self-supporting answer to the value question might be the best we can hope for.

Moreover, even if self-supporting arguments don't seem impressive from the outside, the ability to offer a self-supporting defense is, plausibly, a prerequisite for the viability of any view. So even an ambitious objector should give some weight to the question of whether or not a view can defend itself from the inside. For this objector, then, the arguments here still provide reason to worry about moderate permissivism.

**Conclusion**

Both extreme permissivism and impermissivism take strong positions on how *permissive* rationality is: one is very lenient, and the other is very demanding. Moderate permissivism seemed initially attractive because it strikes a compromise between these two extremes. But in combining the attractive features of extreme permissivism and impermissivism, moderate permissivism seems to lose the benefits of both: it is hard to see why rationality, as moderate permissivism construes it, is especially worth caring about.

If moderate permissivism is right, why does rationality matter? In the absence of a good explanation, we should feel a significant push to one of the extremes: a completely permissive, or a completely impermissive account. Moderation, in this case, is no virtue.
Chapter 3
Immodesty and Educated Guesses

It is a platitude that belief "aims at the truth". Even if you are not sure what exactly this platitude means, it is clear that a belief has succeeded in this aim when it is true, and failed when it is false. What about credences, or partial beliefs? One would hope that these aim at truth as well, in some respect at least. But credences can't be true or false. So what is the relationship between credences and truth? There is no consensus as to what that answer should be.65 One of the central challenges to epistemologists who would like to think in terms of credence, rather than full belief, is to provide an answer to this question.

A second topic, which might at first seem unrelated, is the phenomenon David Lewis calls "Immodesty."66 Think about a magazine like Consumer Reports, which rates consumer products. After gathering information about different products' various features, the magazine makes a recommendation about which product (vacuum cleaner, toaster, etc.) to trust. What if a consumer magazine ranked consumer magazines? If it's to be trusted, Lewis argues, it must at least rank itself as the best. Of course, ranking itself the best doesn't mean that it is the best. Immodesty is a necessary, though insufficient, condition on trustworthiness.

A rational agent, like a good consumer magazine, should be immodest – at least, she should be immodest insofar as she knows that she's rational. Out of all of the possible epistemic methods she could use, she should use the one that she deems best. Another way to put this thought is that a rational agent must (at least) do well by her own lights. Of course, doing well by one's own lights may well be insufficient for rationality. But it is necessary.

An epistemically rational agent must use the method she takes to be best – but "best" in what sense? By this I don't just mean the method she takes to be most rational.

---

65 For discussion of some candidate views, see Hájek [ms]. For reasons of space, I won't get into the details of the proposals he considers.
66 This example comes from Lewis's [1971] paper, "Immodest Inductive Methods". Lewis defines "immodesty" slightly differently – in his terms an "inductive method", rather than the person who follows it, is immodest. (An "inductive method" can be understood as a function from evidence to doxastic states.) I'll follow Gibbard [2008] here in applying the idea of immodesty to an agent who follows an immodest method.
That would be trivial. (And if epistemic permissivism is true, she might even regard several different methods as rational.) Rather, the relevant kind of goodness or choiceworthiness at work here should have something to do with truth. Someone who rationally believes P on the basis of her evidence should regard \textit{believing P} as better than \textit{believing }\neg P \textit{ because, given her evidence, believing P gives her the best shot at believing \textit{truly}. This thought seems obvious, but perhaps not entirely trivial: a rational agent should take her own beliefs to be best in terms of getting to the truth.

If Immodesty holds for rational beliefs, it should hold for rational credences as well. And the right account of Immodesty in a partial-belief framework should work together with the right account of the connection between credences and truth. Giving these two accounts, together, will be the project of this paper.

In the first part of the paper, I'll look at a popular view in formal epistemology, according to which credences get things right by being \textit{accurate}, or close to the truth. Accuracy is measured using a special kind of utility function. This view appears to vindicate Immodesty as well: a rational agent should regard her credences as maximizing expected accuracy according to that utility function. I'll argue that this view does not provide a satisfactory explanation of Immodesty, which casts doubt on its plausibility as the right view of the relationship between credence and truth.

In the second part of the paper, I'll look at how we might explain Immodesty without appealing to epistemic utility functions. If epistemic utility theory doesn't work, can we hold onto the thought that the correct explanation for Immodesty should appeal to truth? Some, like Gibbard, say no. I suggest an alternative view that says yes. Rational credences get things right by licensing true \textit{educated guesses}. And a rational agent should regard her credences as optimal because, from her perspective, they give her the best shot at guessing truly. I will then discuss some further benefits of the guessing picture as an interpretation of the relationship between credences and truth. Most importantly, I will suggest that it can be used to argue for probabilistic coherence as a rational requirement.

1. Epistemic Utility Theory
Since credences can't themselves be true or false, the relationship between credences and truth must be indirect. A current movement in formal epistemology holds that, when it
comes to credences, what we should care about is *closeness* to truth, or accuracy. Just as we strive for truth in our beliefs, we strive for accuracy in our credences. In this section I will look at a way we might try to use this branch of formal epistemology—"epistemic utility theory"—to answer the questions I raised at the beginning. First I will briefly set out the relationship between credences and truth, as conceived by epistemic utility theorists. Then I will look at how epistemic utility theory could be used to explain or vindicate Immodesty. Finally, I will raise some worries about this approach.$^{67}$

1.1 Accuracy and epistemic value

Epistemic utility theorists start with an intuitive notion of accuracy. The accuracy of a degree of credence in P is determined by its closeness to the truth value of P: 1 if P is true, and 0 if P is false. Suppose P is true. Then your credence in P is *accurate* to the extent that it’s close to 1; it’s *inaccurate* to the extent that it’s close to 0. If your credence in P is .8, and mine is .9, mine is more accurate than yours. The opposite holds if P is false. A maximally accurate credence in P is one that matches the truth value of P. So a maximally accurate credence function, covering all of one’s degrees of belief, is one that’s fully opinionated—it assigns either 0 or 1 to everything in its domain—and omniscient.

So far all of this is fairly uncontroversial, but also very rough. Clearly, .9 is closer to 1 than .8 is—but how much more accurate is it? (Should accuracy be measured linearly, or some other way?) Questions like this are answered by an agent’s "scoring rule", a utility function that assigns value to credences depending on their distance from the truth. Different scoring rules might answer these questions differently. According to one simple scoring rule, Absolute Distance, the inaccuracy of a degree of credence depends linearly on its absolute distance from the truth. According to the Brier Score, the inaccuracy of a degree of credence is the square of the absolute distance. It is better to be more accurate, so all rational agents should have a scoring rule according to which the

---

$^{67}$ One caveat before I begin: the way I present the dialectic here is very different from how it has gone in recent literature. That’s because the aim of most epistemic utility theorists is to justify things like probabilistic coherence and conditionalization—Immodesty is not their focus. However, I think that my presentation is not entirely unfaithful to the way epistemic utility theorists think of the constraints on acceptable accounts of epistemic value: Immodesty is often in the background of these debates, even if it is not explicitly discussed. For those reasons, I don’t think that the differences in presentation will be a problem.
utility of one’s credence goes up as it gets closer to the truth. (So both Absolute Distance and the Brier Score, which assign zero inaccuracy to omniscient credences, would be permissible if this were the only restriction.) Beyond that, for all we have said so far, there is a lot of room for variation.

Epistemic utility theory’s key move is to treat accuracy as a special kind of value. This opens the door to a new way of understanding epistemic rationality, on an analogy with decision theory: just as we can model practically rational agents as aiming to maximize utility, we can model epistemically rational agents as aiming to maximize epistemic utility.

What could differences between agents’ scoring rules represent? One way to gloss the issue is to say that different scoring rules represent different weightings of the “Jamesian” goals of believing truth and avoiding error. This story holds that, as James observed, these two goals need to be traded off against one another. Someone who cared only about believing truth would believe everything – contradictions included – so as to guarantee having the most true beliefs possible. Someone who only cared about avoiding error would suspend judgment on everything so as not to risk believing anything falsely. Though neither of these extremes seems rational, there might still be a number of intermediate ways that we could rationally trade off the two. Some scoring rules encode relatively greater concern with believing truth by rewarding credences with higher and higher accuracy as they reach the extremes: the difference between .8 and .9 might count for more than the difference between .6 and .7. These scoring rules “encourage” becoming more opinionated. Other scoring rules encode relatively greater concern with avoiding error by assigning less and less extra expected utility as they reach the extremes. By the lights of these scoring rules, becoming more opinionated is “riskier”.

The potential loss involved in moving from .7 to .8 might not be worth the relatively

---

68 For instance, in this passage from Joyce [2009]: “The more convex a score is at a point, the greater the emphasis it places on the avoidance of error as opposed to the pursuit of truth near that point. The more concave it is, the greater the emphasis it places on the pursuit of truth as opposed to the avoidance of error. As William James famously observed, the requirements to avoid error and to believe the truth – epistemology’s two ‘great commandments’ – are in tension, and different epistemologies might stress one at the expense of the other. [The James/Clifford] debate plays out in the current context as a dispute about convexity/concavity properties of measures of epistemic accuracy.” (p. 281) Though not everyone interprets scoring rules this way, I think the Jamesian story is a prima facie plausible interpretation of what it is that scoring rules represent. I will raise some problems for this Jamesian story later.
small potential increase in epistemic utility. According to this story, an agent’s scoring
rule represents the tradeoff between the Jamesian goals that she, personally, is willing to
make.

1.2 Accuracy and Immodesty
We began with the thought that, with the right theory of the relationship between
credences and truth, we should expect a vindication of Immodesty. And on one popular
approach in epistemic utility theory, we get one. The idea is that, just as practically
rational agents take the action that maximizes expected utility, *epistemically* rational
agents adopt the credences that maximize expected *epistemic* utility, or expected
accuracy. We might put this roughly as a constraint on which combinations of credences
and scoring rules can be rational: an agent’s credences and scoring rule must satisfy the
requirement that her credences maximize expected accuracy by the lights of her
credences and her scoring rule. From a rational agent’s point of view, then, her own
credences are optimal. No other credences have higher expected accuracy than her own.

We need to strengthen this requirement a bit in order to fully vindicate Immodesty. A disanalogy between action and belief is that, in the case of action, there is
nothing wrong with an agent choosing arbitrarily between two equally-good options. If
chocolate and vanilla look equally good, from your point of view, you are free to pick
either one, or choose by flipping a coin. But in the case of belief – especially on a fine-
grained, degreed-belief picture – it seems that whenever there are two equally-good
options, there will also be a third option that strikes a compromise between the two. A
rational agent should (non-arbitrarily) adopt the belief that makes that compromise. There
would be something wrong with an agent who regarded believing P and believing ~P as
equally good, truth-wise, but adopted one or the other belief rather than suspending
judgment. So unlike the decision-theoretic case, it seems insufficient for a rational agent
to merely maximize expected epistemic utility in her credences. She should *uniquely*
maximize expected accuracy.69

---

69 We have to be careful here: of course it’s rationally permissible to regard *some* credence functions as
having higher expected accuracy than one’s own. For example, the credence function that assigns 1 to all
truths and 0 to all falsehoods has higher expected accuracy than my own. But while I can pick out that
credence function under that description, I can’t pick it out by specifying whether it assigns 0 or 1 to each
Finally, we need to make a restriction on the kinds of credences that we want Immodesty to be true of. Immodesty, as we are currently thinking about it, is a rational requirement: rational agents (who know that they are rational, etc.) must be immodest. But this requirement takes “wide scope”. We want our picture of epistemic value to vindicate Immodesty for agents who in fact have rational credences. But it’s not necessary – and in fact, would likely be a mistake – to try to vindicate Immodesty for irrational credences as well. For instance, we might not want the result that someone who has credence 1 in everything, contradictions included, takes herself to maximize expected accuracy. (And in fact, if we are working with expected accuracy, it’s not even clear what this would amount to: expectation is a technical notion that is defined only for probability functions.)

This line of thought suggests one last constraint on our picture of epistemic utility: Immodesty should be supported only for agents with probabilistically coherent credence functions. To sum up, here is what we want our Immodesty principle to say, on the epistemic utility theory picture:

**Accuracy Immodesty (first pass):** Rational agents should take their own credences to uniquely maximize expected accuracy.

Scoring rules that meet this desideratum are called “strictly proper”. If epistemic utility theory is to provide a satisfactory vindication of Immodesty, we must hold that only strictly proper scoring rules are rationally permissible.  

---

70 A common desideratum for epistemic utility theorists like Joyce is that Immodesty should hold for all and only those credence functions that are probabilistically coherent. Why only the probabilistically coherent ones? Because, Joyce argues, incoherent credences face a special problem that is not shared by any coherent credences: they are accuracy-dominated by coherent credence functions. So incoherent credences can be ruled out a priori as rational. (And: we don’t need Immodesty to hold for irrational credences as well as rational ones.) Why should Immodesty hold for all probabilistically coherent credences? Joyce argues that no probabilistically coherent credence function can be ruled out a priori: even for credences that might look crazy, he claims, there is some possible body of evidence that could support having those credences rationally. (See Joyce [1998] and [2009].)

71 One wrinkle here: while much of the epistemic utility theory literature focuses on expected accuracy, some of the arguments (like Joyce’s arguments for probabilistic coherence) use dominance reasoning alone. So not all epistemic utility theorists need to buy into the full picture I discuss here. However, if we want to give an account of Immodesty, we cannot appeal to dominance reasoning alone for the same reason that
It turns out that strict propriety is a hard demand to meet. Not all ways of measuring accuracy support this thought: by the lights of some initially-attractive scoring rules, rational agents will regard a credence function other than their own as maximizing expected accuracy. One example of an *improper* scoring rule is the Absolute Distance rule, which I mentioned above. This is a surprising and informative consequence, according to epistemic utility theory. Though Absolute Distance might have seemed from the outset like a perfectly reasonable way to think about accuracy, it is not. An agent who valued accuracy according to Absolute Distance would feel pressure to make her non-extreme credences extreme; she would not regard her non-extreme credences as the optimal response to her evidence. She would not be immodest. But, plausibly, it can be rational to have non-extreme credences. (The whole point of thinking in terms of credences rather than full beliefs, or full beliefs plus suspension of judgment, is that sometimes these intermediate attitudes *are* appropriate.) So, a rational agent cannot value accuracy according to Absolute Distance.

Nevertheless, there are many scoring rules that *are* strictly proper; there is more than one way to measure accuracy and still end up Immodest. If epistemic utility theory is on the right track, then, it teaches us that strictly proper scoring rules are the only permissible ways to trade off the potential benefit of believing the truth and the risk of believing falsehoods. It also teaches us that some of our pre-theoretic thoughts about accuracy—such as the thought that it could be measured legitimately by Absolute Distance—cannot appeal to (non-uniquely) maximizing expected accuracy alone. Suppose you have probabilistically coherent credences and a strictly proper scoring rule, and know that your credences are not dominated. That’s good—it shows that your credences display the epistemic virtue that you care about. But it’s not enough. Lots of other coherent credence functions are also non-dominated. So from your point of view, your own doxastic state is no better than lots of other doxastic states you could have had instead. Dominance reasoning alone, therefore, cannot fully explain why an ideally rational agent stably holds her own credences rather than switching to others.

Here’s a quick illustration of why Absolute Distance isn’t strictly proper. Suppose you have .6 credence in P, and .4 credence in ¬P (so, your credences are coherent). You are wondering whether .6 credence in P maximizes expected accuracy—or, equivalently, minimizes expected inaccuracy. The inaccuracy of credence .6 is just the absolute distance between .6 and 1 if P is true, and the absolute distance between .6 and 0 if P is false. So to find the expected inaccuracy, take the average of these values, weighted by the likelihood that each of these possibilities is the actual case: (.6)(1-.6) + (.4)(0-.6) = .48. Is .6 the best—i.e., most expectedly-inaccurate—credence in P to have, from your point of view? No: any credence higher than .6 is better. To take the extreme case, consider credence 1 in P. The expected inaccuracy of credence 1 is (.6)(1-1) + (.4)(0-1), or .4. In fact, using the Absolute Distance measure, only credences of 1, 0, and .5 will ever maximize expected accuracy from their own point of view. So unless your credences have only those values, if your scoring rule is Absolute Distance, you will never maximize expected accuracy by your own lights.
Distance – are wrong. In its vindication of Immodesty, epistemic utility theory also tells us something substantive about the relationship between rational credence and truth.73

Here is the final proposal we end up with, if we understand Immodesty in the way suggested by epistemic utility theory:

**Accuracy Immodesty**: Rational agents should take their own credences to maximize expected accuracy, as assessed by a strictly proper scoring rule.

In the next section I’ll raise some worries for this approach.

## 2. Objections to epistemic utility theory

In the previous section I set out a way in which we could use epistemic utility theory to answer the questions we started out with. *Distance* from the truth, or accuracy, is the analog of truth in credences; it is measured by strictly proper scoring rules. And strictly proper scoring rules, along the way, carry Immodesty as a direct consequence.

Should we adopt this picture? I think there are reasons not to – at least insofar as we are trying to explain Immodesty. The main worry is that, though epistemic utility theory claims to take a decision-theoretic approach to epistemology, its parallels with decision theory are quite tenuous. Though it asks us to understand accuracy as a special kind of utility or value, this kind of value does not end up playing a very substantive role in the final epistemological theory. There are also reasons to worry about epistemic utility theory on simplicity grounds: the framework seems to build in extraneous details that have little explanatory payoff.

### 2.1 Which strictly proper scoring rule you have makes no difference

The first reason for skepticism about epistemic utility theory stems from a disanalogy between epistemic utility (as used in epistemic utility theory) and practical utility (as used in decision theory). Take a standard decision-theoretic case: I want to decide whether to steal a bite of your sandwich. I am not sure whether you ordered it with pickles. I don’t like pickles, so I would prefer not to take a bite if there are pickles on the sandwich. But

---

73 Though it’s worth noting that this argument can be tollensed: if what epistemic utility theory says about accuracy is too surprising, one might object that it’s simply not talking about accuracy anymore. Gibbard [2008] and Maher [2003] each raise versions of this worry. Though I take this to be a serious concern, I’ll set it aside here: even if you’re not moved by these objections, there are other reasons to worry about epistemic utility theory.
if there are no pickles, I would prefer to have a bite. Rational action depends on both my
beliefs and my preferences; changing either one might change what is rational for me to
do.

In the practical realm, it is common to take this thought a bit farther: supposing
that there are many different permissible utility functions, we might conclude that there
are many different potential rational responses to a given decision problem. For example:

Sweet or Salty? At the diner, you and I both face the following choice: an ice
cream sundae or a basket of onion rings. I strongly prefer sweet snacks to salty
ones, and you strongly prefer salty snacks to sweet ones. We share our evidence
regarding the sweetness and saltiness of our choices.

Plausibly, our preferences – supposing that both are rationally permissible – justify
different actions in response to our shared set of options. Since I strongly prefer the ice
cream, that’s what I should order; I’d be acting irrationally if I ordered the onion rings.
And the opposite goes for you: you should order the onion rings, and you shouldn’t order
the ice cream. Permissivism about rational preference – that is, the thought that there are
many different rationally permissible preference orderings – therefore seems to support
permissivism about rational action.74

Is the same true in about beliefs? Does your scoring rule make a difference to
what it is rational for you to believe? If we follow the “Jamesian” reasoning from before,
it looks like perhaps it should. Just like Sweet or Salty, it seems like we could get a case
where different epistemic values justify (in a purely epistemic sense) different responses
to a shared body of evidence. This example, adapted from Kelly, has at least prima facie
appeal:

Election: We both have the same evidence regarding the upcoming presidential
election. We agree that on balance it favors D: the hypothesis that the Democrat
will win. But our epistemic values differ: I value believing truth more highly than

74 More carefully, it seems to support “interpersonal” permissivism: there is more than one rational
response to our shared decision problem, but only one rational response for me. One might think that
intrapersonal permissivism is true as well, especially in cases where many options maximize expected
utility (or where many options meet the relevant threshold, for satisficing views). I’ll focus on the
intrapersonal type of permissivism in my comparison with the epistemic case. As I have already argued, we
should not think that many different doxastic states could maximize expected accuracy from one rational
agent’s point of view.
you do, and you value avoiding error more highly than I do. So you and I have different scoring rules.  

If our scoring rules are both permissible, it seems plausible that could be rationally required to have different credences in D in response to our shared evidence. Maybe in this situation, my credence in D should be .7, and yours should be .6. Since I care more about believing truth, it will seem worth the risk to me to become more opinionated; since you care more about avoiding error, you prefer to play it safe. And just like in the practical case, it’s plausible that if your epistemic values changed, what it would be rational for you to believe would change too.

If epistemic utility were just like practical utility, scenarios like Election could justify permissivism about belief. We could argue for permissivism about rational belief on the basis of permissivism about epistemic utility, just like we might argue for permissivism about rational action on the basis of permissivism about preferences or utilities. Or we could argue against permissivism about rational belief by arguing that only one epistemic utility function is rational – just as a hedonistic utilitarian, for example, might argue that there is only right way to weigh the disvalue of a stubbed toe against the value of a hug.

But epistemic utility, on the common picture I’ve been describing, does not play the same role as practical utility. Scenarios like Election cannot come about – or if they can, the part about epistemic values is a red herring. Though epistemic utility theory supports a limited kind of permissivism about rationally acceptable scoring rules, it does not support permissivism about rational credence. Which strictly proper scoring rule you have does not make a difference to what it is rational for you to believe.

To see why, let’s first look back at the overall picture that epistemic utility theory supports. First, good news for the Jamesian: epistemic utility theory is compatible with permissivism about both credences and epistemic value. Strictly proper scoring rules are used to justify the Bayesian formal constraints of coherence and conditionalization, but no further substantive constraints on what one’s initial credences (or “priors”) should

---

This example comes from Kelly [2013]. Kelly puts the point in terms of different epistemic values, rather than scoring rules in particular (though scoring rules seem to me to be a natural way of fleshing out Kelly’s thought more precisely).
look like. As long as you start off with coherent credences, have a strictly proper scoring rule, and update on your evidence by conditionalization, you will maximize expected accuracy by your own lights. This means that, for all we have said so far, epistemic utility theory is compatible with the thought that any coherent priors can be rational. After seeing the sun rise the last hundred days in a row, one rational Bayesian agent might be highly confident on that basis that the sun will rise again tomorrow. Another might be highly confident, on the same basis, that tomorrow a giant tomato will rise instead. And both may count as maximizing expected epistemic utility by their own lights.

As we saw earlier, epistemic utility theory also allows for permissivism about rationally acceptable scoring rules. The only restriction is that scoring rules be strictly proper – they must support Immodesty for probabilistically coherent agents. Though this restriction eliminates a surprisingly large swath of ways to measure accuracy, it still leaves several possibilities open.

So epistemic utility theory leaves us with permissivism about belief and permissivism about value. But unlike in decision theory, permissivism about value in this domain does not justify permissivism about belief. Which strictly proper scoring rule you have does not matter at all. Becoming an impermissivist about rational belief would not justify becoming an impermissivist about epistemic value. And becoming an impermissivist about epistemic value would not justify becoming an impermissivist about rational belief.

Return to Election. Suppose that both of us have coherent credences and update by conditionalization, and that we have different strictly proper scoring rules. (In other words: we are both rational by the lights of epistemic utility theory.) We share all of our evidence relevant to D. My credence in D is .7, and yours is .6. My credence in D is .7, and by the lights of my credences and my strictly proper scoring rule, .7 is the credence in D that uniquely maximizes expected accuracy. What if I evaluate my credence using your scoring rule? Well, your scoring rule is strictly proper, too. So if I use your scoring rule to evaluate my credences, I will still get the result that my credences maximize expected accuracy. And the same is true for you, from your own point of view. Using your credences and your strictly proper scoring rule, your credences uniquely maximize expected accuracy; using your credences and my strictly proper scoring rule, your
credences still uniquely maximize expected accuracy. This is not an accident: it is exactly what we wanted when we imposed strict propriety as a requirement on scoring rules. Strictly proper scoring rules are designed to make any coherent credence function come out as optimal, by its own lights. So switching from one strictly proper scoring rule to another provides no pressure to switch from one coherent belief state to another.

This means that the Jamesian story, as we told it before, cannot be right. Even if we have different scoring rules, encoding different ways of weighing up the goals of believing truth and avoiding error, those differences don’t show up in our credences. Epistemic value, as understood by epistemic utility theory, isn’t what makes the difference.

What does make the difference, then, if rational credence is permissive? The answer is our priors. In fact, if we think through what the Bayesian framework requires, it becomes clear that once that framework is in place, there is no room for scoring rules to make a difference. We could rewrite Election as follows:

Given total evidence $E$, $P_{me}(D) = .7$, and $P_{you}(D) = .6$

Well, since we are both updating by conditionalization, it must have been the case beforehand that:

Prior to receiving evidence $E$, $P_{me}(D|E) = .7$, and $P_{you}(D|E) = .6$

Our current credences are dictated by our prior conditional credences, together with our evidence. We don’t have to mention our scoring rules or epistemic value at all to tell this story.\(^{76}\)

Is it a problem that, on this picture, epistemic value doesn’t make a difference? On the one hand, we shouldn’t have expected it to make a difference in the first place. Stepping back for a moment and looking at full beliefs, this does not seem like a problem at all. When I introduced Immodesty, I claimed that the principle should almost sound trivial: if you believe $P$, then you should regard your belief as optimal because, by your lights, believing $P$ is the best way to have a true belief about $P$. And in a full-belief framework, we don’t have to mention truth at all to describe how you actually go about responding to your evidence; it doesn’t obviously play a substantive, separate role, like

\(^{76}\)But wait, what happened to the Jamesian argument? I will come back to it in section 7.
preferences or values do in practical reasoning. (Why do you now believe Q? Because you believed “If P, then Q”, and you just learned P.) So the analogue of truth, in a degraded-belief framework, should play a similar, almost trivial part in explaining how you should respond to your evidence. Why should we have to mention truth, or accuracy, in explaining how a rational agent should assign her credences?

But on the other hand, if the relation between truth and credences is properly understood as a kind of value, we should expect it to make a difference. The fact that strictly proper scoring rules do not make a difference shows, I think, that they should not be understood as utility functions. The analogy between epistemic rationality and practical rationality is at best very tenuous, and at worst, merely a (not very informative) metaphor. But this analogy is supposed to be a major innovation of epistemic utility theory. In this respect, I think, epistemic utility theory is on the wrong track.

2.2 Your ranking of suboptimal options makes no difference

In the last subsection, I argued that, since epistemic utility behaves very differently from practical utility, it is a mistake to think of it as a kind of utility at all. In this subsection I’ll bring out this point by looking at a different, but related disanalogy between practical and epistemic utility. That is: scoring rules, unlike practical utility functions, have far more structure than is necessary for the work that they are meant to do. In this subsection I will bring out why this extra structure does no work; if that’s right, we have reason to prefer a simpler theory if we can find one.

By “extra structure” I mean this: scoring rules tell us not only which credences are best, but which are second-best, third-best, and so on. We can use scoring rules to assess the accuracy of any credences whatsoever, including those that are incoherent – and we can assign all of these credences both cardinal and ordinal values. In this way, scoring rules are analogous to moral or practical utility functions. For example, a hedonistic utilitarian’s moral value function doesn’t just tell her which action to take; it also ranks her suboptimal options and tells her how well each one does with regards to promoting the balance of pleasure over pain. Or consider this case: you want to get to the pharmacy, which closes in 15 minutes. You have three options: walk, bike, or drive. Balancing the aesthetic and environmental merits of each option against the risk that it will not get you
to the pharmacy on time, suppose the expected utilities of these three options are as follows:

<table>
<thead>
<tr>
<th>Option</th>
<th>Expected Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk</td>
<td>-5</td>
</tr>
<tr>
<td>Bike</td>
<td>0</td>
</tr>
<tr>
<td>Drive</td>
<td>10</td>
</tr>
</tbody>
</table>

Your practical utility function assigns a score not only the best option, or the one that maximizes expected utility, but to the others as well.

These “extra” preferences or values make no difference to this decision considered in isolation (since after all, you just need to pick the option that maximizes expected utility). But in the practical realm, this rich structure can be both useful and explanatory when it comes to predicting and explaining your rational behavior in nearby cases. Suppose that your options change: you get to your car and see that it has a flat tire. In this new decision problem, with driving removed as an option, the relative goodness of walking and biking come into play. You should now choose to ride your bike. Because our circumstances can change, it makes sense for you to assign value to walking and biking, and to rank them against one another, even when driving is obviously best. The comprehensive value assignments supplied by a practical value function have a substantive to play in determining what we ought to do, in a broad sense, even if not every detail is relevant in every case.

Furthermore, the details of a practical value function — including the way in which it rates suboptimal options — can help explain what makes actions right and wrong. By giving a full story of what’s valuable, we can help explain why driving is better than walking and biking, and why biking is best when driving is not an option. The same factors that determined to the value of biking while driving was an option should, ceteris paribus, also determine the value of biking when the only alternative is walking. (It is still environmentally friendly, a good way to get exercise, etc.) This is especially true for simple consequentialist moral or practical theories, where goodness can be straightforwardly aggregated: the reason one action is right, rather than another, is that it produces more hedons or utiles. More nuanced consequentialist theories might be able to make the same claim. A utility function represents how different kinds of moral reasons
weigh against one another. The reasons that count in favor of biking rather than walking, when driving is an option, should generally be a good indication of which reasons count in favor of biking when driving is not an option.

These considerations arguably support the use and relevance of detailed, comprehensive value functions in the practical realm. But in the epistemic case, because of Immodesty, a rational agent’s opinions about suboptimal options never come into play. If a rational agent has no new evidence, the credences that maximize expected accuracy for her are always the ones she already has. And if she does have new evidence, the credences that maximize expected accuracy are always her prior credences conditionalized on that new evidence—no matter which credences were second- or third-best before receiving that new evidence. Supposing that epistemic rationality is a matter of responding correctly to your evidence, those two cases exhaust the possibilities that matter for determining your beliefs. There are no situations where we can predict a rational agent’s new rational credences by looking at her ranking of suboptimal options in nearby cases.

We might try to cook up cases where these opinions do matter by “removing” some of your doxastic options, just as we did in the practical case. (After all, in the practical case, your suboptimal opinions mattered because your options for acting could be more or less restricted.) In fact, this is precisely the type of case imagined in Moss [2011]: she considers cases where an agent has to change her credences to compromise in response to disagreement, or must choose between actions that will, because of non-rational persuasion or brainwashing, result in different belief states. In these cases, as Moss conceives them, simply responding rationally to your evidence is not an option: you must choose between credence functions that do not maximize expected accuracy from your current perspective.

Moss argues that, in these scenarios, we should adopt the credences that, among our remaining options, maximize expected accuracy. This might seem like a sensible rule, given the analogy with moral or practical cases: by choosing the most (expectedly) accurate credence available to you, you are simply making the best of a bad situation. But Moss’s scenarios are importantly different from moral or practical choices where all of your options are bad ones. That’s because an agent need not act immorally—in fact, she
need not even depart from ideal morality – if she chooses the best of a set of bad options. It might well be possible to act in a morally ideal manner even if, for example, your actions will produce a great deal of suffering. (An epistemic analogue of this might be a case in which you don’t have much evidence. In that case, you might realize that the credences that maximize expected accuracy for you won’t be very accurate. But in adopting those credences, you can maintain ideal rationality.) Instead, Moss’s examples are ones in which all of your options are irrational.

The closest moral analogy to Moss’s cases might be a moral dilemma: a situation in which all of your options are not just bad, but also immoral. You might be skeptical that moral dilemmas exist, or that the right moral theory should give a definite answer about what to do when faced with one. But even setting those worries aside, it’s at least implausible that our theory of morality should hinge on features of moral decision that are only present in moral dilemmas. Likewise, it’s implausible that our theory of epistemic rationality should be hostage to features that are only relevant in rational dilemmas. If anything, we should get an account that works in normal cases where everything is going as it should, and figure out what to say about rational dilemmas afterwards. So even if Moss is right about what to believe in these dilemma cases (and maybe she is, if epistemic utility theory is the right approach) we shouldn’t take these cases as our basis for positing such rich structure in scoring rules, or for adopting epistemic utility theory in the first place.77

The details of your epistemic value function – how you rank suboptimal epistemic options – have no predictive or explanatory role. Can we do without them? Insofar as we

---

77 Pérez Carballo [ms] suggests another way in which scoring rules might influence what we should believe: by differentiating among different lines of inquiry we might pursue to gain credences in propositions about which we previously had no opinion. This is similar to Moss’s strategy in that epistemic utility only makes a difference when our options are limited. (In Pérez Carballo’s case, we start off with limited options; the epistemic change we undergo involves expanding the domain of our credence function.) Pérez Carballo’s examples may be more promising than Moss’s as a justification of the use of scoring rules: they do not as obviously involve irrationality, so they are not as obviously epistemic dilemmas. Even so, the cases are arguably still far enough from paradigm cases of rational belief change that we should not hinge the rest of our epistemological theory on how to resolve them. This topic deserves much more discussion, but I will not pursue it more here.
prefer simpler theories, we should look for a theory that doesn’t require us to ascribe these extraneous preferences or values.

2.3 Summing up

Epistemic utility theory, especially when we accept a restriction to strictly proper scoring rules, does not give us an adequate explanation for Immodesty. Although its main distinguishing feature is a special kind of value, the role of value turns out to be very limited in determining what we should believe. While epistemic utility theory started out as a way to make the Jamesian goals precise, it can’t do justice to the Jamesian thought that different epistemic values might justify different beliefs. And the theory posits a rich and detailed structure to our values that turns out to be explanatorily inert. For those reasons, I doubt that this project gives us much explanatory insight into the attitudes rational agents should have towards their own beliefs, and the connection between rational credence and truth.\(^{78}\)

The factor that really determines what we should believe isn’t epistemic value, but our priors. And priors aren’t values; they’re just more beliefs.\(^{79}\)

3. Immodesty without epistemic value

\(^{78}\) This isn’t to say that epistemic utility theory needs to be rejected wholesale. Even if it fails to explain the phenomena at issue here, it might still be useful for other purposes. We might compare epistemic utility theory, in this respect, to a “consequentialized” Kantian or rights-based theory in ethics. Several ethicists have argued that it’s possible to express non-consequentialist theories in consequentialist terms with the help of an appropriately nuanced utility function. Kantians have a legitimate complaint with this approach: while a utility function might get the right results, they might claim that it’s not the right way to understand morality, and doesn’t give a satisfactory explanation for the moral facts. But consequentialized Kantian theories might be very useful in other domains: formal semantics, for example. Similarly: strictly proper scoring rules like the Brier score, with their complex ranking of suboptimal credal states, do not help us explain Immodesty. But they may still be good for other purposes.

\(^{79}\) More carefully: what determines our beliefs isn’t epistemic value if epistemic utility theory is the right way to think about epistemic value. Might there be some other understanding of epistemic value, or another way to use the machinery of epistemic utility theory, which has some bearing on our choice of priors? One thing to note is that the factors that made epistemic utility theory unsatisfying in this context are quite general: that it satisfies Immodesty, and that you need beliefs or credences to make use of its notion of value. This new way of understanding epistemic value would have to be one that could be used by an agent independently of her beliefs (if it is to be used to choose her priors). And to fully vindicate the Jamesian line of thought, the new way of understanding value should not respect Immodesty. If we find Immodesty independently plausible, we might worry about any conception of epistemic value that makes sense of the Jamesian thought. (I’ll come back to this in section 7.) We might also have reason to worry about the whole idea of “choosing priors” out of nowhere. These are interesting questions, but I won’t have the space to pursue them fully here. Thanks to Steve Yablo for suggesting this line of response.
I began by looking for a way to explain Immodesty. In the last section I looked at one attempt to do so, using epistemic utility theory; that seemed not to work.

Maybe there is a way to explain Immodesty without positing epistemic value. But how? One way is to appeal to a different kind of value instead. Gibbard, for example, proposes an explanation for Immodesty that appeals to a rational agent’s *practical* values. I’ll look at that account in this section, and raise some problems with it. Then I’ll suggest an alternative account that (I hope) does better – one that doesn’t appeal to either epistemic or practical value, and which can be adopted in both consequentialist and nonconsequentialist frameworks.

Gibbard argues that we can explain Immodesty by looking to a rational agent’s practical interests. His claim is this: if a rational agent compares her own credences to others with respect to which actions they recommend, she should conclude that acting on her own credences will maximize expected value. So a rational agent should regard her own credences as best for the purposes of maximizing expected practical value.

Gibbard’s argument, partially adapted from Schervish, goes roughly as follows. Suppose you are choosing an “action policy”, which includes a credence function and a utility function. Whenever you need to decide what to do, your action policy will recommend the option that maximizes expected utility as assessed by that credence function and utility function. Holding your actual utility function fixed, which credence function would you now choose to have as part of your action policy?

If you are rational, Gibbard argues, you should prefer to act on your own credences. That’s because your current credence function is the only one that, no matter which options you face, is guaranteed to recommend the action that you now regard as preferable. The easy way to state this would be: your credence function is the only one such that maximizing expected utility according to that credence function guarantees maximizing expected utility from your point of view in every situation. If you chose a different credence function to act on, there would be some situations in which that
credence function (plus your action policy) recommended doing something that, by your current lights, looks like the wrong decision.\footnote{Gibbard proves this by imagining a continuum of bets on a proposition, $S$, at odds $1-x:x$, where $0<x<1$ (so, all possible odds at which you could bet on $S$ are represented). You will win $S(1-x)$ if $S$ is true, and lose $Sx$ if $S$ is false. Your action policy will recommend accepting bets that have positive expected value, and rejecting bets that have negative expected value. In other words, if “$Pr$” is the credence function you will use for acting, your action policy will recommend accepting bets where $Pr(S) < x$. You will reject bets where $Pr(S) > x$. If you act on some credences other than your own, you will end up accepting bets that, by your own lights, should be rejected, and vice versa.}

We might put Gibbard’s proposal like this:

**Action Immodesty:** A rational agent should take her own credences to be best, by her current lights, for the purposes of *taking the right actions*.

Gibbard’s solution offers a simple and elegant justification of Immodesty: it doesn’t require us to posit any special kind of value, or to give up our intuitive grasp of “accuracy”. (Since rational beliefs aren’t understood as maximizing expected accuracy, it’s perfectly fine on this account to endorse an improper scoring rule – like Absolute Distance – according to which your own beliefs may not maximize expected accuracy.) But there are several reasons to think that we should try to do better. Gibbard’s explanation of Immodesty takes what we thought of as a purely epistemic phenomenon and turns it into a practical one; epistemic rationality depends entirely on our practical concerns. This seems like changing the subject. Furthermore, we might think that even an agent with *no* practical interests or capacity to act might still be subject to the requirements of epistemic rationality. Gibbard’s account rules this out.

More plainly, we might still think that a good explanation for Immodesty should appeal primarily to the importance of truth, and should explicate a sense in credences “aim at the truth” for its own sake. The intuitive thought that we started off with, which gave truth a central role, has all but disappeared on Gibbard’s account. Gibbard is well aware of this: he bills his account as an alternative to what he sees as a failed project of trying to explain Immodesty on truth-related grounds alone. But we should not throw in the towel just yet. We should hold out for a better explanation of Immodesty: one on which a rational agent should prefer to keep her credences not because they will get her what she wants, but because she takes them to represent the world as it is. In the next section I will propose such an explanation.
4. The guessing account

I’d like to suggest the following interpretation of Immodesty:

**Guess Immodesty**: A rational agent should take her own credences to be best, by her current lights, for the purposes of *making true educated guesses*.

I will argue that the guesses warranted by your own credence function have the best expected success; so, you should choose your own credences for the purpose of making true guesses. This proposal ties Immodesty to truth, but does so without requiring agents to have a utility function (epistemic or otherwise), and without requiring agents to have unnecessarily detailed or comprehensive opinions about which credence functions are second- and third-best.

This account requires us to introduce something new: the notion of an “educated guess”. We might also just call it a potential forced choice, supposing that your sole concern in choosing is to reach the truth. Suppose you are given two or more options—say, P and ¬P—and have to choose between them. All you care about is answering truly. If you are rational, your guess will align with the option you’re most confident of. If you are more confident of P than of ¬P, you should guess P. If you are more confident of ¬P than of P, you should guess ¬P. The educated guesses I’ll be talking about here are the guesses warranted by your credences.

Gibbard’s picture asked us to see epistemically rational agents as actors, preparing to make good decisions no matter what options they might face. The guessing picture asks us to see epistemically rational agents like students preparing to take a multiple choice test. Even if you aren’t sure of the right answers—after all, you don’t know everything—you should, insofar as you want to get the answers right, take your best shot.

To see how Immodesty follows from this picture, consider the following hypothetical scenario. You will take an exam. The exam will consist of just one question regarding some proposition that you have considered (and in which you have some degree of credence). You will have to give a categorical answer—for example, “It will rain tomorrow”—and you will not have the option of refusing to answer. For the purposes of this exam, you only care about answering truly. Now suppose that, for the

---

81 This strategy is similar to Gibbard’s hypothetical series of bets, borrowed from Schervish. (It might be helpful to think of my general line of argument as a “depragmatized” version of Gibbard’s.)
purposes of taking this exam, you must choose an *educated guessing policy*. It consists of a credence function, plus the norm described above: when faced with a forced choice between two or more propositions, your educated guess should be the proposition in which your credence is highest. When faced with a forced choice between two or more propositions given some supposition, your educated guess should be the proposition in which your conditional credence (conditional on the supposition’s being true) is the highest. My claim is that if you are rational, then the prospectively best guessing policy, by your current lights, will include your current credences.

For concreteness, let’s call your *current* (rational) credence function “Cr”, and let’s call the credence function you *should* pick for the purposes of guessing “Pr”. My claim, then, is that Pr = Cr. You should pick your own credences as the right credences to use for guessing.

Here is a warm-up example. Suppose the exam question is the following:

**Q1: Is it raining?**

In this case, if Pr(Rain) > .5, you should guess “yes”. If Pr(~Rain) > .5, you should guess “no”. (What if Pr(Rain) = Pr(~Rain)? Let’s say that in that case you should decide in some arbitrary manner.)

To be prepared to answer questions like Q1, you don’t need to pick your own credences. All you need is that Pr has the same “valences” as Cr – it assigns values that, for every proposition that it ranges over, are on the same side of .5 as the values that Cr assigns. (If Cr assigns .6 to Rain, and Pr assigns .8, for instance, Pr will recommend the answer that is right by your current lights.) In fact, to answer questions like Q1, you might as well choose the credences recommended by “Absolute Distance” – that is, the maximally-opinionated counterpart of Cr. So Q1 doesn’t yet show us that Pr = Cr.

But not all questions are like Q1. For example, consider the following, slightly more complicated question:

**Q2: Supposing that it’s not sunny, which is it: rain or snow?**

Suppose you’re highly confident that it is sunny – but if it’s not sunny, it’s more likely to be rainy than snowy. Your current credence in Rain is .2 and your current credence in
Snow is .05. For this question, you will not do well if you choose the credences recommended by Absolute Distance. (If you choose the credences recommended by Absolute Distance, you will have credence 1 that it will be sunny, and credence 0 in both rain and snow. Those credences will recommend answering Q2 by flipping a coin or answering arbitrarily. But, by your current lights, this is a bad idea; it’s better to guess that it will rain. That will give you the best shot at answering truly, which is what you care about.) To answer questions like Q2, you don’t just need to choose a credence function with the same “valences” as yours – you also need one that differentiates among different possibilities as you do, including those that are unlikely to come about.

Choosing a guessing policy that will do well on questions like Q2 still doesn’t get us to the conclusion that Pr = Cr. (You might choose Pr so that it is uniformly a bit more or a bit less extreme than Cr, and still get the same recommended answers to questions like Q2.)

Now consider another, more complicated question:

**Q3:** Suppose all of the following: a weighted coin has “Rain” written on one side, and “~Rain” on the other. It is weighted 1-x:x in favor of whichever of Rain or ~Rain is true, where 0 < x < 1. Now suppose:

(a) the coin is flipped, out of sight;
(b) you answer whether Rain; and
(c) you and the coin disagree about Rain.

Who is right?

Suppose you’re more confident of Rain than of ~Rain. Just flip the argument around if the opposite is true for you. What this question is doing is eliciting some of your conditional credences: specifically, your credence that the coin will be right and your credence that you will be right, supposing that you disagree. Intuitively, you should take yourself to be more reliable than the coin – more likely to get the right answer – if your credence in Rain is higher than the weighting of the coin. You should take yourself to be less reliable than the coin if your credence in Rain is lower than the weighting of the coin. So if you disagree, you should guess that you were right iff Cr(Rain) > x. You should guess that the coin was right iff Cr(Rain) < x.

Suppose you know that you will guess Rain (because Cr(Rain) > Cr(~Rain)). Call your credence in Rain y and the coin’s weighting x. What we are looking for is this:
conditional on the coin saying \(-\text{Rain}\), what’s the probability that the coin is right? What’s the probability that you are right?

In general, the probability that the coin says \(-\text{Rain}\) will be the sum of the probability that it says \text{Rain} and is right, and the probability that it says \text{rain} and it’s wrong. Or: \((1-y)(x) + (y)(1-x)\). The following therefore gives you your conditional credences:

\[
\begin{align*}
\text{Cr(Coin is right|Coin says } \neg\text{Rain}) &= (1-y)(x) / ((1-y)(x) + (y)(1-x)) \\
&= (x - xy) / ((1-y)(x) + (y)(1-x)) \\
\text{Cr(Coin is wrong|Coin says } \neg\text{Rain}) &= (y)(1-x) / ((1-y)(x) + (y)(1-x)) \\
&= (y - xy) / ((1-y)(x) + (y)(1-x))
\end{align*}
\]

To see which of the conditional credences will be higher, just look at the numerators (the denominators are of course the same). The quantity \(xy\) is the same in both cases, too. So it’s easy to see that if \(x > y\), the first conditional credence will be greater than the second; if \(y > x\), the second will be greater than the first. So you should guess that the coin is right, conditional on your disagreeing, if your credence in \text{Rain} is greater than the weighting of the coin. You should guess that you are right, conditional on your disagreeing, if your credence in \text{Rain} is less than the weighting of the coin.

This means that in order to get the best guessing policy, you should want to choose \(Pr\) such that, for any proposition \(R\), \(Pr(\text{Rain})\) is “on the right side” of \(x\), for any value that \(x\) could take. So \(Pr = Cr\). Any credence function other than \(Cr\) will recommend wrong answers for some propositions and some values of \(x\) – in response to questions like Q3, it will recommend guesses that you presently regard as more likely to be false than true. So if you care about answering each question truly, you should choose \(Cr\) for your guessing policy.

5. Vindicating probabilism
The explanations for Immodesty have, so far, assumed that rational agents must have probabilistically coherent credences. But we might also want an argument for probabilism. Epistemic utility theory gives an answer in the form of accuracy-dominance arguments; the guidance or action approach gives an answer in the form of Dutch Book
arguments. I want to suggest here that the guessing account can also be used to argue for probabilism.

Probabilism is traditionally expressed in three axioms. There are different formulations; here are the three I’ll use. Assuming that $P$ is any rational credence function, $T$ is a tautology, and $Q$ and $R$ are disjoint propositions:

\[
\begin{align*}
\text{Non-Triviality:} & \quad P(\neg T) < P(T) \\
\text{Boundedness:} & \quad P(\neg T) \leq P(Q) \leq P(T) \\
\text{Finite Additivity:} & \quad P(Q \lor R) = P(Q) + P(R)
\end{align*}
\]

My strategy here will be to show that if you violate Non-Triviality or Boundedness, you will guess falsely in situations where guessing falsely is not necessary. It is irrational to guarantee unnecessary false guesses. Therefore it is irrational to violate these axioms. I will then give a different kind of argument for Finite Additivity.

By “unnecessary guaranteed false guesses”, I mean guaranteed false guesses in situations where, were you to change some of your credences, you would not be guaranteed to guess falsely. For example, if your credence in a necessary falsehood is greater than zero, there will be situations in which, when choosing between the necessary falsehood and some contingent alternative (to which you also grant nonzero credence), you are licensed to guess that the necessary falsehood is true. This is an unnecessary guaranteed false guess: if your credence in the necessary falsehood were zero, you would be licensed to guess that the contingent proposition was true, and your guess would not be guaranteed false. (This is similar to part of the argument for Boundedness, below.)

Compare this to a situation where you are choosing between two necessary falsehoods. Even if your credence in each necessary falsehood is zero, you are licensed to guess that one or the other is true (because you have no other choice). This bad outcome shouldn’t be held against you. There is nothing you could do to avoid it: no matter what credences you adopt in the two necessary falsehoods, you would not be able to avoid making a guaranteed-false guess.

### 5.1 Non-Triviality

\[
\begin{align*}
\text{Non-Triviality:} & \quad P(\neg T) < P(T)
\end{align*}
\]
Non-Triviality says that your credence in a tautology, T, must be greater than your credence in its negation, ~T. We can prove this axiom in two parts. First suppose that your credence in ~T is greater than your credence in T. This immediately leads to problems: if you were to guess whether T is true or false, you would be licensed to guess that T is false. But T is a tautology, and therefore guaranteed to be true. So your guess is guaranteed to be false. And it is unnecessarily guaranteed false: if your credence in T were greater than your credence in ~T, your guess would be guaranteed true.

Second, suppose that your credence in ~T is equal to your credence in T. If you were to guess whether T or ~T, you would be licensed to guess either way. This means that you would be licensed to guess ~T, which is guaranteed to be false. This guess is also unnecessarily guaranteed false: if your credence in T were greater than your credence in ~T, you would not be licensed to guess ~T in this situation, so you would not be licensed to make a guaranteed-false guess.

5.2 Boundedness

\[
\text{Boundedness: } P(\neg T) \leq P(Q) \leq P(T)
\]

Boundedness says that it is irrational for you to be more confident of some contingent proposition than you are of a necessary truth, and it is irrational for you to be less confident of a contingent proposition than you are of the negation of a necessary falsehood. One way to read this axiom is as saying that, of all of the possible credences someone could have, your credence in necessary truths must be the highest and your credence in necessary falsehoods must be the lowest. If we add in a plausible assumption about what this means, we can prove Boundedness within the educated guess framework.

Let’s look at how things go with the assumption first. The assumption is this: there is a maximal (highest possible) degree of credence, and a minimal (lowest possible) degree of credence. I’ll also assume a plausible consequence of this assumption in the guessing framework. First: if you have the maximal degree of credence in some proposition, A, you are always licensed to guess that A when A is one of your choices. (You may also be licensed to guess in favor of a different choice if your choices are A and some other proposition in which you have the same credence.) In a question like Q3, above, you are licensed to guess that you were right about A for any possible weighting.
of the coin. Second: if you have the minimal degree of credence in some proposition, B, you are never licensed to guess B (except in cases where you are licensed to guess arbitrarily between B and another proposition in which you have the same, minimal, degree of credence). In a situation like Q3, you are licensed to guess that the coin was right about B for any possible weighting of the coin.

For simplicity, let’s assume that your credences satisfy Non-Triviality, which we have already argued for. So, your credence in T is greater than your credence in ~T.

Assuming that there is a maximum credence and a minimum credence, we can normalize any agent’s credences, assigning the maximum credence 1 and the minimum credence 0. So, if your credence in T is maximal, it should be equal to 1. If your credence in ~T is minimal, it should be equal to 0.

First, let’s prove that your credence in T should be maximal. Suppose that your credence in T is not maximal: it is 1 - y, where 0 < y < 1. Now consider the following question, similar to Q3, above:

**Q4:** A weighted coin has T on one side, and ~T on the other. It is weighted 1-x:x in favor of whichever of T or ~T is true, where 0 < x < y. Now suppose:

(a) the coin is flipped, out of sight;
(b) you guess whether T; and
(c) you and the coin disagree about T.

Who is right?

To get the situation clear: the coin is weighted in favor of whichever of T or ~T is true, that is, it is weighted in favor of T. You will guess T (because you satisfy Non-Triviality). So if you and the coin were to disagree, the coin would be guaranteed to be wrong. But given the setup of Q4, and your credence in T, you should guess that, were you to disagree, the coin would be right: you should regard the coin as more reliable than your credence when it comes to guessing whether T. Your non-maximal credence in T causes you to get Q4 wrong. So you should not have a non-maximal credence in T.

Now let’s prove that your credence in ~T should be minimal. Suppose your credence in ~T is z, where 0 < z < 1, so z is not the minimal credence. Consider the following question:

**Q5:** A weighted coin has some contingent proposition R on one side, and ~R on the other. It is weighted 1-x:x against whichever of R or ~R is true, where
0 < x < z. Now suppose that the coin is flipped out of sight. Your question is: which is right? The coin (however it landed), or ¬T?

In this case, the coin is very unreliable; it is weighted against the truth about R. Your credence in ¬T is also very low. So from your perspective, both options ¬T, and whatever the coin says about R – are unlikely to be true. Out of those two options, your credences license you to choose ¬T. However, ¬T is guaranteed to be false; so, your guess is guaranteed to be wrong. And it’s unnecessarily guaranteed false: if your credence in ¬T were minimal, you would never be licensed to guess in favor of the coin, no matter how it is weighted. So you should not have non-minimal credence in ¬T.

What we’ve shown is that you should have maximal credence in T and minimal credence in ¬T. There is no possible coin that you should take to be more reliable than your credence in T, and no possible coin that you should take to be less reliable than your credence in ¬T.\(^8\)

5.3 Finite Additivity

Finite Additivity: \[ P(Q \lor R) = P(Q) + P(R) \]

While Non-Triviality and Boundedness provide constraints on our credences in necessary truths and falsehoods, Additivity is about how our credences in contingent propositions should fit together with one another. Contingent propositions are not themselves guaranteed to be true or false. So violating Additivity – while it will lead to some irrational guesses – may not lead to any guaranteed false guesses. That means that we need a different argument for accepting Additivity as a rational constraint. I will provide such an argument, and then address a potential objection.

\(^8\) Traditional Dutch Book Arguments seem to prove something stronger: that not only should your credence in T be 1, and your credence in ¬T be 0, but also that your credence in Q should not be higher than 1 or lower than 0. Having such credences is interpreted as licensing certain (bizarre and irrational) betting behavior; since this behavior is not rational, the credences that license it are not rational either. My argument does not go this far. But I don’t think this should be a concern in the guess framework: there is no way to guess that Q more than 100% of the time, or less than 0% of the time. Since there is no such thing as having a credence of more than 1 or less than 0, there is no additional reason to argue that these credences would be irrational to hold.
Suppose your credence in Q is x, your credence in R is y, and your credence in Q \lor R is z. Finite Additivity says if you are rational, x + y = z. Intuitively, my argument is going to bring out the fact that, if you violate Additivity, you take your rational credence in (Q \lor R) to be more (or less) reliable than the sum of your credence in Q and your credence in R. (Or maybe more intuitively: you take your total evidence to be a better (or worse) indicator of (Q \lor R) taken as a disjunction than of Q and of R, taken individually.) The way that you guess regarding Q and R will differ depending on how the options are presented to you.

As before, the argument for Additivity is broken into two cases. First, suppose z > x + y. Now compare the following two guessing scenarios:

**Q6a:** A certain weighted coin, Coin 1, has “yes” on one side, and “no” on the other. It is weighted 1-v:v, where x + y < v < z, in favor of “yes” if and only if (Q \lor R) is true. Now suppose that the coin is flipped out of sight, and you guess whether (Q \lor R). You are to say “yes” if you guess (Q \lor R), and “no” if you guess \neg(Q \lor R). Your question is: if you and the coin disagree, which is right?

In this scenario, you take your guess to be more reliable than the coin. You are guessing whether (Q \lor R), and the coin is less opinionated than you are on the matter. So you should guess that, supposing you disagree with the coin, you are right. And this is the only guess that your credences warrant in this case.

**Q6b:** Coin 1 has “yes” on one side, and “no” on the other. It is weighted 1-v:v, where x + y < v < z, in favor of “yes” if and only if (Q \lor R) is true. There are also two other coins. Coin 2 has Q on both sides, and Coin 3 has R on both sides. Now suppose that all three coins are flipped out of sight, and you guess “yes” or “no” in response to this question: Did at least one of Coin 2 and Coin 3 land true-side-up? Suppose you and Coin 1 disagree. Who is right?

In this scenario, you take your guess to be less reliable than the coin. You are guessing about Q and R, taken separately, and the coin is more opinionated than you are. So you should guess that, if you disagree, the coin is right. And this is the only guess that your credences warrant in this case.

This combination of guesses reveals an inconsistency in your credences. In Q6a, you preferred your own credences over the coin, and in Q6b, you preferred the very same
coin over your credences. But the only difference between Q6a and Q6b was in how you guessed: directly on (Q v R) in Q6a, and on both Q and R, taken separately, in Q6b.

We can create a parallel setup for the case where z < x + y. (Coin 1' will have to be given a weight, w', such that z < w < x + y.) Again, you will guess inconsistently: you will take the coin to be more reliable than you are when you guess directly regarding (Q v R), and you will take the coin to be less reliable than you are when you guess regarding both Q and R, taken separately.

This is irrational. The only factor bearing on both the reliability of Coin 1 and the truth of your guesses in Q6a and Q6b is whether at least one of Q or R is true. So you have no basis for treating the two cases differently. But if you violate Additivity, your credences require you to treat the two differently.

Here is another way we could put the point. Your guess in Q6a reflects your judgment about the strength of your evidence regarding Q and R. Guessing in favor of yourself, over the coin, makes sense because you consider your evidence to be a stronger indicator of whether Q or R is true than the coin is. From the perspective of your evidence, trusting the coin over your own guess is a positively bad idea; it gives you a worse shot at being right. Compare this to your guess in Q6b. This guess also reflects your judgment about the strength of your evidence regarding Q and R in comparison to the coin. From the perspective of your evidence, as characterized in Q6b, trusting your own guess over the coin is a positively bad idea. But the relevant evidence – the evidence bearing on Q, and the evidence bearing on R – is the same, and you are judging its strength in comparison to the very same coin. So it doesn’t make sense for your credences to license both guesses simultaneously. The only rational option is to obey Additivity.

I’d like to close by addressing a worry you might have about this argument. That is: you might think that providing a different kind of argument for Additivity from the kind we had for Boundedness and Non-Triviality is a weakness of the guessing picture. After all, the existing justifications of probabilism, Dutch Book arguments and epistemic utility theory, defend the three axioms in a unified way. The Dutch Book argument says that agents with incoherent credences will be licensed to take bets that guarantee a net
loss of money (or utility, or whatever you’re betting on). Epistemic utility theorists argue
that incoherent credences are accuracy-dominated by coherent credences, or else that
incoherent credences fail to maximize expected epistemic utility. In each case, the same
type of argument is given in support of each axiom.

I’d like to propose that we think of Additivity differently. The argument I gave
was meant to show how, if your credences violate Additivity, you will fail to make sense
by your own lights. How reliable you take yourself to be regarding Q and R depends on
how you are asked about Q and R – how the very same guessing situation is presented to
you. This is the same sort of argument we might make to show that it is irrational to
consider John a bachelor but deny that he is unmarried. You’re not guaranteed to have
any particular false belief about John by virtue of believing both that he’s a bachelor and
that he’s married. But you will have beliefs that don’t make sense by your own lights (at
least if you know what bachelors and marriage are). We could also make this kind of
argument to show that it is irrational to believe both P and my evidence supports ~P. In
both cases, we might not have the security of a formal dominance-style argument on our
side. But that doesn’t show that the rational requirements in question don’t hold.

Of course, my argument for Additivity depends on some controversial
assumptions. Most obviously, I relied on the thought that the evidence bearing on (Q v R)
was the very same evidence as that bearing on both Q and R, separately. Some would
deny this: for example, they might argue that it’s possible to have evidence bearing on (Q
v R) as a disjunction that has no bearing on either Q or R individually. (Alternative
formal systems, like Dempster-Schafer theory, deny Additivity as a rational requirement,
and are supported by these kinds of motivations. My argument would do little to persuade
advocates of Dempster-Schafer theory to become Bayesians. So you might think this
shows that the guessing account can’t really provide a strong justification of probabilism.

I take this to count in favor of the guessing account. It can be used to make sense
of, and argue for, the axioms of probability for those who are sympathetic to certain
background assumptions. But I think it is also flexible enough that, were we to deny these
assumptions, we might still be able to make use of the general framework. The guessing
picture leaves room for some of the substantive debates in formal epistemology. And the
particular argument I proposed for Additivity makes clear where the substantive
assumptions come in. A more ambitious project might attempt to give a stronger justification for particular proposed rational requirements, working within the guessing framework. I won’t attempt to take that project on here, but will instead leave this as a first step in the direction of what such a project might look like.

6. Some more advantages of the guessing account

We’ve seen how the guessing account can vindicate Immodesty. In this section I’ll mention some advantages of the account, both as an alternative to the epistemic utility theory account and as a project that is independently useful and worth pursuing. I’ll also point to some avenues for future applications of the account.

First of all, the guessing account doesn’t commit us to a theory of “epistemic value”, or ask us to think about epistemic rationality in decision-theoretic terms. In order to be immodest, according to the guessing account, a rational agent just needs the following attitude: that for any question that she might face, she aims to answer it truly. This requirement says nothing about how she might weigh the importance of believing truly about various propositions; she need not have opinions at all about which belief state is second-best, third-best, and so on. (Of course, we can add on a separate theory about this if we want to.) We are free to think of believing truly about P and believing truly about Q as incommensurable values. This means that we can respect what Selim Berker calls the “separateness of propositions”: we need not buy into a “consequentialist” account of epistemic rationality. 83

Second, the guessing account leaves us free to accept any account of accuracy that we want. Scoring rules aren’t required to vindicate Immodesty, but we might want them for some other purpose. If we like the Absolute Distance rule, or some other scoring rule that is not strictly proper, we can use that as our measure of accuracy without worrying about fitting it into a decision-theoretic picture of epistemic rationality.

83 A non-consequentialist guessing picture could also provide an argument for why trade-offs between our beliefs are not rational: remember that if you adopt credences other than your own for your guessing policy, your guessing policy will recommend particular guesses that are likely to be false by your current lights. On a non-consequentialist interpretation, we might say that adopting this policy is irrational because it fails to respect the truth of that particular proposition. See Berker [2013a] and [2013b], as well as Carr [ms] for discussion of epistemic tradeoffs.
These advantages provide some initial motivation for taking the guessing picture seriously. Of course, taking on this account also raises a challenge: we need to accept *educated guesses* as a new piece of theoretical machinery. What are educated guesses? Why should we think they exist, let alone think that they are something worth caring about?

There are two ways to argue in favor of including educated guesses in our overall epistemological theory. One is to find instances of them “in the wild”. I think this strategy has some promise. Multiple choice tests are one way to elicit educated guesses in the sense I’m interested in here. Assertion under time constraints might be another way to elicit them, as well as unconstrained assertion. Statements like “if I had to guess...[P] but I’m not sure...” plausibly express educated guesses in the same sense.

But while that first strategy might well pan out, I think we don’t have to rely on it. That’s because we have an alternative strategy at our disposal – to argue for the existence of educated guesses by putting them to theoretical use. The argument here, which uses educated guesses to explain Immodesty, is one part of that project. Using educated guesses to argue for probabilism is another part. And I think there are other places, too, where educated guesses could be put to work. They may be especially helpful in cases where reliability is important. Reliability is typically thought of in all-out terms, as a matter of one’s propensity to get things right and wrong. But since credences cannot be true or false, it’s not clear how they can get things right and wrong in the first place. It might be helpful, then, to tie credences to something that can be directly assessed for truth or falsity: something like educated guesses. For example, consider “higher-order” evidence: evidence about what your evidence is, or what it supports. Plausibly, doubts about your own rationality, or about your ability to assess or identify your evidence, should rationally affect what you believe. This is because rationality normally goes along with reliability; irrational belief is often unreliable belief as well. The inclination to speak in all-out terms when discussing the import of higher-order evidence is clear in much of the literature on this topic, but the guessing picture makes it explicit. 84 I have also argued

84 For example, see White [2009]'s “Calibration Rule”, which states: “If I draw the conclusion that p on the basis of any evidence E, my credence in P should equal my prior expected reliability with respect to P.” See also Elga [2007]'s (similar) formulation of the Equal Weight View: “Upon finding out that an advisor disagrees, your probability that you are right should equal your prior conditional probability that you would
elsewhere for an explicit connection between credences and educated guesses, in the interest of spelling out how higher-order evidence works.\textsuperscript{85} If this strategy is successful, it will be another mark in favor of including educated guesses as a legitimate part of our epistemological toolbox.

7. Lingering question: what happened to the Jamesian argument?

The second challenge we face involves the Jamesian goals: while they were a large motivating factor for epistemic utility theorists, they don’t seem to make an appearance at all on the judgment account. What happened to them?

There is some initial plausibility to the thought that that our epistemic goals – in particular, the Jamesian goals – should influence what we believe. \textit{Election}, the case I discussed earlier, is actually one that Kelly introduces to argue for exactly this point.\textsuperscript{86} Field appeals to a similar thought:

\begin{quote}
[W]e recognize that a slight modification of our goals – an increase in the relative value of reliability over power – would lead to a preference for [another] system, and we regard the alternative goals as well within the bounds of acceptability. Consequently we make no very strong claims for the preferability of our system over the alternative: the alternative is slightly less good than ours given our precise goals, but slightly better on alternative goals that are by no means beyond the pale.
\end{quote}

The idea here is that having different acceptable “goals” can rationalize different ways of responding to evidence. Field characterizes these goals slightly differently from Kelly (or James), referring to “reliability” and “power” rather than “believing truth” and “avoiding error”. But the idea seems to be similar. A more “reliable” belief-forming method is one that avoids taking a stand, and so doesn’t often go wrong. A more “powerful” belief-forming method is one that risks making more mistakes, but also gets things right more definitively.

If Kelly and Field – and maybe James himself – are right, permissivism about epistemic values should support permissivism about rational belief. But, as I have argued, epistemic values do not play this role when they are understood in the way suggested by

\textsuperscript{85} See Horowitz and Sliwa [ms].
\textsuperscript{86} See Kelly [2013].
epistemic utility theory. The key to this argument was Immodesty: since any rational way of trading off between the Jamesian goals required a rational agent to see herself as making the best possible such tradeoff, I argued that switching from one rational set of epistemic values to another didn’t bring with it any pressure to change one’s beliefs. So this argument applies to more than just epistemic utility theory. It should apply to any account that has Immodesty as a consequence.

So what happened to the Jamesian goals? What role, if any, should they play? Are they distinctly epistemic values, and do they make a difference to rational belief?

As I see it, there are three main options for what we could say here. The first is to reject the thought that the Jamesian goals are epistemic values at all. If we do have preferences between believing truth and avoiding error, they don’t rationally influence what we should rationally believe. This strategy leaves room for the thought that Jamesian values might have practical import; we might set up a situation where the \textit{practical} rewards for having a true belief were high, and there was no penalty for believing falsely; in that situation it might be practically rational to try to make oneself more opinionated. But, according to this strategy, it would be a mistake to think that situations like \textit{Election}, where our concerns are purely epistemic, could work the same way. This strategy seems plausible to me on its own, but it fails to do justice to Kelly’s original Jamesian argument.

A second option is to cash out the Jamesian goals as a kind of epistemic value, but deny that they rationally affect how we should respond to our evidence. Instead, they might affect things like evidence-gathering behavior: someone who cares a lot about having true beliefs (as opposed to avoiding error) might be interested in gathering more evidence, so that she could rationally form stronger opinions about the world. Though this strategy might seem closer to the spirit of the Jamesian thought, it also incurs some possibly-unwelcome commitments: insofar as the Jamesian goals are relevant to rational action, they will be in competition with practical goals like eating lunch or catching the bus on time. Depending on how we understand these conflicts, this second approach therefore threatens to blur the boundary between epistemic and practical rationality.

Finally, we could accept the idea that the Jamesian goals are distinctly epistemic values, and insist that they do play a role in determining rational responses to our
evidence. This idea seems to me to have the least plausibility of the three. To hold this position, we need to give up on some other prima facie compelling thoughts about rationality: for instance, that it requires coherence, or that it requires Immodesty for agents who know that they are rational.

In sum, I think that we should reject the Jamesian line of thought, at least if it purports to describe a purely epistemic phenomenon. For any interpretation of the Jamesian values that respects Immodesty, the Jamesian argument for permissivism fails. There may well be different permissible epistemic values, and there may well be different permissible responses to a body of evidence – but the former thesis can’t justify the latter, at least not in a purely epistemic way. A broader conclusion may be warranted as well: that the whole idea of epistemic value, as an analogue to moral or practical value, is misplaced.

**Conclusion**

We began by observing that rational agents should not want to change their credences without receiving new evidence: they should be “immodest”, or happy with the credences they have. Why is this? One prominent line of thought explains this phenomenon by positing accuracy measures as a special kind of epistemic value. But once we make certain plausible restrictions about what’s rational, there is no work left for value to play in determining what we should believe.

The best way to explain Immodesty, then, may not involve epistemic value at all. I’ve proposed one way we might go: seeing credences as conducive to truth, not accuracy, and assessing this truth-conduciveness by looking at the judgments that our credences recommend.
Bibliography


Carr, Jennifer. [ms] “How to Expect When You’re Expecting”.


– [2011] "Disagreement, Question-Begging and Epistemic Self-Criticism", *Philosopher’s Imprint* 1, no. 6


– [ms] “Lucky to be Rational.”


Hájek, Alan. [ms] “A Puzzle About Partial Belief”.


Horowitz, Sophie and Sliwa, Paulina. [ms] “Respecting All The Evidence”.


Meacham, Christopher. [ms] “Impermissive Bayesianism”.


Pérez Carballo, Alejandro. [ms] “Good Questions”. Available at: http://perezcarballo.org/files/good.pdf


Schoenfield, Miriam. "Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief". Forthcoming in *Noûs*.


Titelbaum, Michael. [ms] “In Defense of Right Reasons”.


Weatherson, Brian. [ms] “Do Judgments Screen Evidence?”


– [notes a] “Improbable Knowing”.

– [notes b] “Very Improbable Knowing”.

96