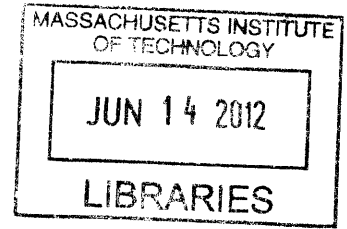


Epistemic Levels

by

Daniel Greco

A.B., Princeton University (2006)
M.Phil., Cambridge University (2007)



ARCHIVES

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

©2012 Daniel Greco. All rights reserved.

Author

Department of Linguistics and Philosophy
May 1, 2012

Certified by

Roger White
Associate Professor
Thesis Supervisor

Accepted by

Roger White
Chair of the Committee on Graduate Students

The author hereby grants to MIT permission to reproduce and to
distribute publicly paper and electronic copies of this thesis document
in whole or in part in any medium now known or hereafter created.

Epistemic Levels

by

Daniel Greco

Submitted to the Department of Linguistics and Philosophy
on May 1, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In this dissertation I defend some controversial “level-bridging” principles in epistemology. In the first chapter, I defend the KK principle—the principle that if one knows that P , then one knows that one knows that P . I argue that saying plausible things about higher-order *interpersonal* knowledge requires rejecting some popular arguments against KK, and that many apparent counterexamples to KK can be explained away by appeal to contextualism about knowledge. In the second chapter, I take up the topic of epistemic akrasia—the state of believing some proposition P , while believing that one oughtn’t believe that P . While many take for granted that epistemic akrasia is always irrational (and least implicitly endorse a level-bridging principle in the process), there are some apparently powerful arguments for holding that epistemic akrasia must sometimes be rational. I argue that once we get clearer on the descriptive/psychological question of what sort of state epistemic akrasia *is*, we can resolve this puzzle in favor of the intuitive view that epistemic akrasia is a species of irrationality. In the third chapter, I appeal to level-bridging principles to respond to some recent arguments to the effect that certain epistemological debates are somehow non-substantive, or merely verbal. If my argument succeeds, this constitutes a kind of indirect support for level-bridging principles—if we think that epistemological debates typically *are* substantive, we face some pressure to adopt level-bridging principles to explain this.

Thesis Supervisor: Roger White
Title: Associate Professor

Acknowledgments

I owe thanks to far more people than I mention here; the following is a very partial list of those who have helped me in writing this dissertation.

For insightful comments and helpful discussions, I'd like to thank Brian Hedden, Sophie Horowitz, Justin Khoo, Alejandro Pérez Carballo, Tristram McPherson, Susanna Rinard, Damien Rochford, Miriam Schoenfield, Alex Silk, and Paulina Sliwa. I'd also like to thank audiences at MIT—in particular, the MITing of the minds conference, as well as various meetings of MATTI, Eminees, and ERG—and the ANU for useful feedback. Thanks also to the job market preparation seminar at MIT—especially Stephen Yablo.

For her unwavering support and friendship, thanks to my wife, Ivana Deyrup Greco.

Lastly, I would of course like to thank my dissertation supervisor, Roger White, the two additional members of my dissertation committee, Agustín Rayo and Robert Stalnaker, and Jonathan Vogel, who was a dissertation committee member of mine in all but name. Each of you provided invaluable substantive advice on the content of the thesis, as well as necessary moral support and encouragement. I am extremely grateful.

Contents

1	Could KK Be OK?	11
1.1	Introduction	11
1.2	Two Arguments and an Example	13
1.2.1	The Independent Check Argument	13
1.2.2	The Safety Argument	14
1.2.3	An Example	16
1.3	Knowledge, Information, and Normal Conditions	20
1.3.1	Normal Conditions and KK	20
1.4	KK, Closure, and Context	26
1.4.1	Defending the Parallel	28
1.4.2	Context and Higher-Order Knowledge	29
1.5	Conclusions	33
2	A Puzzle About Epistemic Akrasia	35
2.1	The Puzzle	35
2.1.1	Thesis	35
2.1.2	Antithesis	37
2.2	What is Epistemic Akrasia?	39
2.3	Normative Upshots	46
2.4	Higher-Order Akrasia	49
2.5	Conclusion	52
3	Verbal Debates in Epistemology	55
3.1	Introduction	55
3.2	Two Challenges	57
3.2.1	Alston on Justification	57
3.2.2	The Experimental Critique	60
3.3	Vindicating Epistemological Debates	62
3.3.1	Transmission of Substantiveness	63
3.3.2	Applying (TS) to Epistemological Questions	65
3.3.3	Internalism and Externalism	69
3.3.4	Knowledge	71
3.4	Epistemic Expressivism	74

Introduction

It's common practice in epistemology to distinguish between claims at different epistemic levels. Just as we can ask whether I know that P , we can ask whether I know that I know that P , or whether I know that I know that I know that P , and so on all the way up. Likewise with other epistemological concepts, such as probability, evidence, and rationality. We can ask whether P is probable, but also whether it is probable *that* P is probable, or whether it's rational to believe that P , but also whether it's rational to believe *that* it's rational to believe that P .

Questions about the relationships between claims at different epistemic levels turn out to be relevant to a wide range of epistemological debates. Certain skeptical arguments rely on strong claims about the links between different epistemic levels, and many philosophers have thought that we must reject these level-bridging claims if we are to avoid skepticism. More recently, debates about how we ought to respond to evidence of our own fallibility—for example, debates about the epistemic significance of disagreement—have often hinged on whether some level-bridging claim is true.

Are there any principles connecting what's true at one level to what's true at another? It used to be thought that there were, but the story of the last forty or so years of epistemology has been, to a significant extent, a story of the rejection of level-bridging principles. Opposition to level-bridging principles began with reliability theories of knowledge in the 1970s, continued with William Alston's influential paper "Level Confusions in Epistemology" in 1980, and is still going strong today, in no small part due to the work of Timothy Williamson. In this dissertation, I argue against this trend, and in favor of some controversial level-bridging principles in epistemology.

In the first chapter, I present a qualified defense of the KK principle—the principle that knowing that P entails knowing that one knows that P . In everyday scenarios, most of our interest in higher-order knowledge concerns higher-order *interpersonal* knowledge (a subject's knowledge of what other subjects know) rather than higher-order *intrapersonal* knowledge (a subject's knowledge of what she herself knows). If Alice knows that she loves Bob, we're usually not so interested in whether Alice knows that she knows that she loves him. Rather, we're more interested in whether Alice knows that *Bob* knows that she loves him. Ultimately, I argue that saying plausible things about higher-order *interpersonal* knowledge—in particular, about what it takes to get it—requires rejecting some popular arguments against the KK principle. I do try, however, to accommodate the intuitions of those who have rejected the KK principle on broadly reliabilist grounds. I argue that if we accept a contextualist theory of knowledge, we can explain away many of the apparent counterexamples to KK,

in much the same way that earlier contextualists have argued that we can explain away counterexamples to epistemic closure principles.

In the second chapter, I introduce a puzzle about “epistemic akrasia”—the state of believing that P , while at the same time believing that one oughtn’t believe that P . Many writers simply take for granted that epistemic akrasia is always irrational, and at least implicitly endorse a level-bridging principle in the process: namely, the principle that if you ought to believe that you ought to believe that P , then you ought to believe that P . But there are some apparently powerful arguments for holding that epistemic akrasia must sometimes be rational. I argue that we can get clearer on the normative question of whether epistemic akrasia might sometimes be rational by first asking the descriptive/psychological question of just what sort of mental state epistemic akrasia is. Ultimately, I argue that on one attractive answer to this descriptive/psychological question, (one that dovetails nicely with broadly expressivist approaches to metaepistemology) epistemic akrasia turns out to be very closely related to outright contradictory belief. While this doesn’t completely settle the normative question, it does lend some support to those who take the intuitive position that epistemic akrasia must be irrational, and the level-bridging principle that goes naturally with that position.

In the third chapter, I appeal to level-bridging principles to respond to some recent arguments to the effect that certain epistemological debates are somehow non-substantive, or merely verbal. If the inference from “the evidence supports the claim that P ” to “ P ” is a reasonable one, then (given a further, plausible principle that I state and defend) epistemological questions about whether the evidence supports P will be substantive whenever the question of whether P is. Such inferences are only guaranteed to be reasonable, however, in the presence of level-bridging principles. This constitutes a kind of indirect support for level-bridging principles—if we think that epistemological debates typically *are* substantive, we face pressure to adopt level-bridging principles to explain this.

Chapter 1

Could KK Be OK?

1.1 Introduction

The KK principle—the principle that knowing entails knowing that one knows—has seen better days. In the past it was defended by Plato, Aristotle, Augustine, Averroës, Aquinas, Spinoza, Schopenhauer, and Prichard, to name a few.¹ Today it languishes in disrepute.² Louise Antony refers to it as “the principle, roundly rejected by epistemologists of almost every stripe, that in order to know that *P*, I must know that I know that *P*.” (2004, p.12) Its fall is often thought of as closely linked to the rise of externalist theories of knowledge—theories that de-emphasize the role of the knower’s reasons or justification for her beliefs (at least when reasons and justification are understood along traditional lines), and stress the importance of causal and/or nomological connections between the knower and the fact known.³

Despite its current unpopularity among epistemologists, the KK principle—or at least the failure of popular objections to the KK principle—is often presupposed in work *outside* of epistemology. Much work in philosophy of language, game theory, and even computer science appeals to the idea of “common knowledge.”⁴ The proposition that *P* is common knowledge if not only does everybody know that *P*, but everybody knows that everybody knows that *P*, and everybody knows that everybody knows that everybody knows that *P*, and so on *ad infinitum*. But standard arguments against KK typically have the consequence that common knowledge is unattainable.⁵ If these arguments are sound, we must reject explanations of linguistic, economic,

¹Hintikka (1962, pp.107-8), a staunch defender of the KK principle, mentions these and other prominent historical proponents.

²See esp. Williamson (2000), along with Alston (1980), Sorensen (1988), and others.

³See e.g., Hemp (2006).

⁴David Lewis’ (2002) work on convention is perhaps the most influential example in philosophy. See Vanderschraaf and Sillari (2001) for a helpful survey article on common knowledge that discusses, *inter alia*, uses of the concept of common knowledge outside of philosophy. It also provides some examples of how KK is often assumed in discussions of common knowledge.

⁵I’ll explain why these arguments rule out the possibility of common knowledge in §2. Williamson (2000), who argues against KK, embraces this result. He argues that the unattainability of common knowledge can be used to explain some otherwise puzzling phenomena. In my opinion these phenomena are best explained using other strategies.

and other social phenomena that appeal to common knowledge.⁶ The debate over KK, then, has consequences that extend well beyond epistemology.

In this chapter, I aim to push back against the turning tide of conventional wisdom. Among epistemologists, both defenders and opponents of KK often assume that higher-order knowledge is a very exalted epistemic state. Knowing that one knows, or worse, knowing that one knows that one knows...*ad infinitum*, is agreed by both camps to be difficult or impossible for us mere mortals to achieve. Opponents of KK hold that we can have first-order knowledge without meeting the onerous requirements for higher-order knowledge, while defenders of KK hold that even first-order knowledge requires meeting those strict requirements; this is why the KK principle is often associated with skepticism.⁷ The positive picture I'll present—one fully compatible with the externalism commonly thought to be in conflict with KK—rejects the assumption often accepted by both camps; the picture that I'll develop is one in which acquiring knowledge of arbitrarily high orders is no more difficult than we pre-theoretically take the acquisition of first-order knowledge to be. I motivate this picture by arguing that certain now-standard objections to KK deliver unhappy results when applied to examples in which higher-order *interpersonal* knowledge—knowing what *other* people know—is the focus. I go on to argue that the picture I present provides a more attractive account of such cases.

In §2 I briefly describe two counterarguments to the KK principle. I then provide an example in which it is intuitively plausible that agents can gain higher-order interpersonal knowledge in certain ways. I argue that we cannot simultaneously accept (1) intuitive claims about what it would take for the agents in the example to gain higher-order interpersonal knowledge, (2) a closure principle concerning knowledge, and (3) the two counterarguments to the KK principle discussed at the beginning of §2.

I hope the example will have enough intuitive force on its own to suggest that something is wrong with standard arguments against KK. However, rather than simply resting on the intuitive force of the example, in §3 I present a simple theory of knowledge on which closure and KK hold. I argue that it provides an attractive analysis of the example from §2, and that this reinforces the dialectical significance of that example. In §4, I argue that the same apparent counterexamples to KK that motivate its rejection are paralleled by apparent counterexamples to epistemic closure principles, and the same contextualist maneuvers that have been invoked to defend closure from counterexample apply equally well in the case of KK.

⁶Hawthorne and Magidor (2009) argue that failures of KK scuttle Stalnaker's account of assertion.

⁷Barry Stroud (1984) endorses KK, and appeals to it in offering certain skeptical arguments. Jonathan Adler (1981) also holds that the KK thesis leads to skepticism.

1.2 Two Arguments and an Example

1.2.1 The Independent Check Argument

What does it take to know by vision that some object x is red? Plausibly, (at least if we are externalists of some sort) I must have a visual system that is capable of distinguishing red objects from non-red objects, and that system must function appropriately in some particular case to deliver me the information that some particular object is red.

Assuming I know by vision that x is red, what would it take for me to know that I know that x is red? According to what I'll call the "independent check argument," I would need to avail myself of some *other* belief-forming process that could distinguish scenarios in which I have a reliable color-vision system from scenarios in which I don't. This could involve my having access to the testimony of an ophthalmologist who would have told me if my color vision were bad, or my having memories of situations in which I successfully exercised my color vision (memories that would've been different if my color vision were unreliable). Generally, knowing that P requires using some process that will reliably cause one to believe P only when P is true. Knowing that one knows that P requires having some *further* process that can reliably distinguish between situations in which the process that produced one's belief that P is reliable, and situations in which it is not. Having higher-order knowledge requires that I be able to perform an independent check on whether my lower-order belief was reliably produced.⁸

Exactly how this argument works will depend on how we understand what it takes for a belief-forming process to be reliable, and what it takes for such a process to be able to distinguish between cases in which other processes are reliable and ones in which they are not.⁹ However we flesh out the details, there will be cases in which something like the independent check argument seems to capture why we're happy to attribute knowledge to a subject, but not to attribute higher-order knowledge to that subject. For example, we might hold that a reliable chicken-sexer who has no information about her reliability can know whether a chicken is male or female, but cannot know *that* she knows whether the chicken is male or female, since she has no way of independently corroborating that her chicken-sexing faculty is reliable.¹⁰

⁸Many externalist/reliabilist denials of KK seem to be motivated by something like the independent check argument. It's natural to attribute a version of this argument to Alston (1980), and Antony (2004) also seems to have in mind an argument along these lines.

⁹This sort of argument is very naturally captured in a "sensitivity"-based framework of the sort defended by Nozick (1981). For criticisms of that framework, see Williamson (2000, chap. 7), and Vogel (2007). I don't assume, however, that the independent check argument can *only* be motivated if we assume a sensitivity constraint on knowledge.

¹⁰I don't know where chicken-sexing was first mentioned in an epistemological context, but Gasking (1962) offers one early discussion.

1.2.2 The Safety Argument

A second argument against the KK principle appeals to the claim that if you know that P , then there's no nearby possibility in which you are wrong about whether P .¹¹ The claim is easy to motivate—it's plausible that if you could've easily been wrong about whether P , then you don't know that P .¹² If we can gloss "You could've easily been wrong about whether P " as "there is a nearby possibility in which you're wrong about whether P ," then we get our claim that you can't know that P if there is a nearby possibility in which you're wrong about whether P .

On this picture, a necessary condition for knowing that one knows that P is that there is no nearby possibility in which there is a nearby possibility in which one is wrong about whether P . The argument that this leads to failures of KK turns on the idea that the relevant nearness relation is not transitive.

A spatial analogy will help make this clear. There might be no nearby houses in which serial killers live, but there might be nearby houses in which there are nearby houses in which serial killers live (if, e.g., houses count as nearby to one another iff they are within five blocks, and my house is 7 blocks away from the closest house in which a serial killer lives). Similarly, there might be no nearby possibilities in I'm wrong about whether P , but there might be nearby possibilities in which there are nearby possibilities in which I'm wrong about whether P . If things work out this way, then I may know that P while failing to know that I know that P . We might say that both cases involve being safe, but not safely safe. I'll call this argument against KK the "safety argument."

There are many other arguments against the KK principle. For instance, it seems possible to know that P without believing that one knows. If knowledge requires belief, then this entails that it is possible to know that P without knowing that one knows.¹³ One way this might happen would involve creatures that don't even have the concept of knowledge—couldn't such creatures know things without even being *capable* of knowing that they know things?¹⁴ Those sympathetic to the KK principle might try to meet these points head on.¹⁵ But another strategy involves conceding the points and holding that there are still interesting versions of the KK principle that can survive this concession. For instance, one might retreat to the position that if someone knows that P *and* believes that she knows that P , then she knows that she knows that P , or that if someone knows that P , then she is in some sense *in a*

¹¹See Williamson (2000).

¹²This is a bit too quick; Nozick (1981, p. 193) considers a case in which Jesse James is robbing a bank and his mask slips. A bystander sees James' face and thereby comes to know that Jesse James is robbing the bank, even though she could've very easily falsely believed otherwise (had the mask not slipped). It's cases like these that lead Nozick to opt for a more complicated version of the safety requirement, one that concerns the methods by which one forms beliefs. See footnote 41 for some brief discussion, and (Williamson, 2000, chap. 7) for some extended discussion.

¹³The claim that knowledge requires belief ought to be more controversial than it is. See Myers-Schulz and Schwitzgebel (Forthcoming) for some criticisms. If these criticisms succeed, then there's one fewer obstacle to the truth of the KK thesis.

¹⁴Feldman (1981) discusses this objection.

¹⁵On the possibility of having higher-order knowledge while lacking the concept of knowledge, see the essays from the final section of Stalnaker (1999).

position to know that she knows that *P*.¹⁶

While I'm sympathetic to a hard line response, I won't attempt to rebut objections to KK that turn on the possibility of knowing without believing that one knows. For this reason, my defense of KK will be limited in scope. But it won't thereby be deprived of interest. Part of what's interesting about the independent check and safety arguments is that they seem to threaten *any* interesting version of the KK principle—even the qualified versions that allow for the possibility of knowing without believing that one knows. In the next subsection, I'll present an example that puts pressure on the independent check and safety arguments.

Before getting into the details of my example, however, I want to emphasize some straightforward reasons to be suspicious of the independent check and safety arguments. The arguments are clearly onto *something*. The distinction between knowing, *and* knowing while also having an independent check on the reliability of one's source of information, is a real one and may very well be of some theoretical significance. Likewise with the distinction of being safe from error within a small sphere of nearby possibilities, and being safe from error within a larger sphere. But there aren't strong pre-theoretical reasons to identify *these* distinctions with the distinction between knowing and knowing that one knows.

For one, it's not very plausible that in real conversations, questions about whether people know that they know are as easy to interpret as the anti-KK arguments would suggest; we don't hear questions about higher-order intrapersonal knowledge (as opposed to first-order knowledge) as straightforward questions concerning whether subjects satisfy extra high epistemic standards. It's no coincidence that when cataloging the various challenges facing the intelligence community, Donald Rumsfeld discussed the known knowns, the known unknowns, and the unknown unknowns, but failed to mention the unknown knowns.¹⁷ Had he included such a category, I doubt that people would have had a clear idea of what he had in mind.¹⁸ Cases like that of the unwitting but reliable chicken sexer (in which questions about higher-order intrapersonal knowledge seem to make sense) are *unusual*.¹⁹ In real life, questions like "Sure she knows what time the movie starts, but does she know that she knows what time the movie starts?" would typically be met with puzzlement, rather than seamless accommodation. (If you don't believe me, try asking similar questions of non-philosophers.) While the defense of KK I'll ultimately present is perfectly capable of drawing the distinctions drawn by the independent check and safety theorists, it refuses to identify these distinctions as distinctions between lower and higher-order knowledge. This refusal not only has strong theoretical motivation—e.g., in making

¹⁶For example, Shoemaker (1994) argues that believing that *P* and being fully rational entails believing that one believes that *P*. If we're sympathetic to this argument, and sympathetic to KK, we might endorse a similarly qualified version of KK according to which it holds at least for all fully rational agents.

¹⁷The exact quote is as follows: "[T]here are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we don't know we don't know."

¹⁸To be fair, even without including such a category, many commentators found his remarks to be less than a model of clarity.

¹⁹And as I argue in §4, they can be accommodated by the defender of KK.

room for common knowledge—but as the above considerations suggest, it fits with linguistic usage as well.²⁰

1.2.3 An Example

Jones and Smith are playing a game of poker. The cards have been dealt and Jones has the stronger hand. Call the proposition that Jones has the stronger hand “*P*.” As it turns out, the deal was videotaped by a reliable camera, and the tape clearly indicates who has which cards. Without looking at her cards Jones manages to procure a copy of the videotape, and she watches it.²¹ Though Jones has no information about the reliability of the camera, she implicitly trusts it, and so comes to truly believe that *P*.²² In this context, because coming to believe what one sees on the videotape is a reliable method for forming beliefs, it’s plausible to say that Jones not only believes but *knows* that *P*.

The saga continues. Jones was taped while she was watching the tape. Smith sees this second tape—the tape of Jones watching the original tape. Like Jones, Smith implicitly trusts that the camera is reliable. So when Smith sees the tape of Jones watching the original tape, Smith comes to believe that Jones knows that she has the better hand.²³ If people generally do trust what’s on videotapes (at least in contexts like this), and tapes really are reliable, then this seems like a reliable enough way of coming to believe that Jones knows that *P*; we should allow that Smith *knows* that Jones knows that *P*.

Let’s go through one more iteration. Smith was taped when watching the tape of Jones. Jones sees the tape of Smith watching the tape of Jones watching the original tape, and so comes to believe (and by the same argument as above, know) that Smith knows that Jones knows that *P*. So at the end of all this tape watching, Jones knows that Smith knows that Jones knows that *P*.²⁴

²⁰See also Marušić (Forthcoming), who stops short of defending KK, but does argue that it fits with key aspects of linguistic usage.

²¹I stipulate that she doesn’t look at her cards because it will be important later that she doesn’t have any independent checks on the reliability of the tape—if she looked at her cards, she’d be able to compare what she sees directly with what appears on the tape.

²²In realistic cases, subjects have *some* information about the reliability of videocameras—most of us know that they tend to work properly. However, for reasons that will come out towards the end of this section, I don’t think that this unrealistic stipulation will affect the force of the example—I’d still be able to make my argument using a version of the case in which Jones has a more realistic amount of information concerning the reliability of videocameras, though the argument would be slightly more complicated.

²³This may seem a bit too quick—in order for Smith to form this belief, she doesn’t just have to believe that the camera is reliable, but she must also believe that Jones believes this. Ultimately, this won’t be important—we could set up the case so that both parties have excellent evidence that people always believe what they see on videotapes, or we might add that in each case in which somebody is taped watching a tape, they exhibit a sign of recognition which reliably indicates that they trust what’s on the tape.

²⁴Perhaps all we can really say is that Jones knows that Smith *knew* that Jones *knew* that *P*. But we can set up the case to make it plausible that if Jones knows that Smith knew that Jones knew that *P*, then Jones knows that Smith knows that Jones knows that *P*.

We have a three stage story. At the end of the first stage, by watching a videotape, Jones knows that P . At the end of the second stage, by watching another videotape, Smith knows that Jones knows that P . At the end of the third stage, by watching yet another tape from the same camera, Jones knows that Smith knows that Jones knows that P . Write “ $K_J P$ ” for “Jones knows that P ” and “ $K_S P$ ” for “Smith knows that P .”

Let’s now assume a closure principle on knowledge, to the effect that if S knows that P , and P entails Q , then S knows that Q .²⁵ At the end of the three-stage process, $K_J K_S K_J P$. Knowledge is factive, so if Smith knows that Jones knows that P , then Jones knows that P . So it follows from the content of what Jones knows— $K_S K_J P$ —that Jones knows that P — $K_J P$. So given our closure principle on knowledge, we get the result that $K_J K_J P$; Jones knows that Jones knows that P .^{26,27}

In the remainder of this section, I’ll argue that our current conclusion—that Jones knows that Jones knows that P at the end of our three-stage story—is hard to reconcile with the independent check and safety arguments.

If we admit that Jones knows that Jones knows that P at the end of our story, then we have two options. Either (1) Jones knows that P (but not *that* she knows that P) after the first stage, and acquires new higher-order knowledge after the third stage, or (2) Jones already knows that she knows that P after the first stage of the story. I’ll argue that neither option is attractive for defenders of the independent check and safety arguments.

Why isn’t option (1) attractive for the defender of the independent check argument? Because watching the second tape doesn’t give Jones an independent check on the reliability of the camera. Of course, we might endorse (1) and hold onto the letter of the independent check argument by individuating belief-forming processes very finely. Instead of treating trusting the tape as a single belief-forming process,

²⁵More realistic closure principles might not assume logical omniscience, and could be phrased in terms of being in a position to know that Q when one knows that P and P entails Q , rather than actually knowing that Q in such circumstances. This distinction won’t matter in what follows.

²⁶Kripke (Forthcoming) also argues that considerations involving interpersonal knowledge are relevant to questions about the requirements for higher-order intrapersonal knowledge. Roughly, he argues that knowing that you know that P can’t be *that* hard, since it’s often quite easy to know whether *other* people know that P , or for other people to know that you know that P .

²⁷This might already seem odd, for a reason that the following modified version of the example brings out. Suppose Jones watches the original tape, and then watches a tape of herself watching the original tape. It is very plausible that Jones hasn’t acquired any new higher-order knowledge by watching the tape of herself watching the original tape. To say that this *was* a way of gaining higher-order knowledge would be to endorse illegitimate bootstrapping, of the sort criticized by Vogel (2000), Elga (2007), and others. If watching the tape doesn’t provide new higher-order knowledge in the *intrapersonal* case, why should it do so in the *interpersonal* case? While I haven’t yet introduced the materials necessary to fully answer this question, I can give a preview. If we accept the KK principle, we needn’t say that there’s any bootstrapping going on when Jones watches a tape of herself watching the original tape—watching the second tape won’t be a way for Jones to *come to know* that she knows that P , since she’ll *already* have known that after watching the first tape. But accepting KK doesn’t force us to say that once Jones watches the original tape, she also knows that *Smith* knows that P , or that Smith knows that *Jones* knows that P . So if KK holds, Jones’ watching successive tapes can give her new higher-order *interpersonal* knowledge, but not new higher-order *intrapersonal* knowledge, since there’s no new higher-order *intrapersonal* knowledge to be had.

we might hold that for each tape, trusting what's on *that* tape constitutes a new, independent process. Then watching later tapes could provide independent checks on earlier tapes, and Jones could start out with only first-order knowledge, but acquire higher-order knowledge by watching another tape from the same camera (thereby getting an independent check on the content of the first tape). In the absence of a solution to the generality problem, we can't rule out this strategy on principled grounds.²⁸ But it does seem to me that individuating processes at such fine a grain is contrary to the spirit of the independent check argument.

An example may help this point sink in. Suppose that you are in fact trustworthy, though I have no way of telling this. You tell me that *P*. The spirit of the independent check argument would suggest that I can know that *P* on the basis of your testimony, but cannot know that I (or you) know that *P*, since I lack an independent check concerning your trustworthiness. Suppose you go on to tell me that you're trustworthy—or to make the case more precisely analogous to the example already under discussion, you tell me that somebody else knows that you know that *P*. It doesn't seem as if your making this further claim should let me know that I (or you) know that *P*, if I didn't already know that you were trustworthy—if I don't know whether or not you're trustworthy, your telling me that you are can't dispel my ignorance, at least *ceteris paribus*. The case involving the camera seems relevantly analogous—if I don't start out knowing that the camera is reliable, it's plausible that seeing more tapes from the camera in which the camera seems to be getting things right can't dispel my ignorance. So, if the independent check requirement is to do the work it's meant to, such further tapes shouldn't count as providing independent checks on the reliability of the camera.

Option (1) is just as unattractive for the safety theorist as it is for the defender of the independent check argument. There doesn't seem to be any sense in which watching the successive tapes is strengthening Jones' epistemic position with respect to *P* itself. The only information Jones has that bears on whether *P* is the original tape, and other tapes of people watching that original tape. It doesn't seem as if possibilities in which the camera malfunctioned and misrecorded the players' hands are any more remote once Jones has watched tapes of people watching the original tape (or tapes of people watching tapes of people watching the original tape) than they were when Jones had only watched the original tape. While Jones seems to end up with higher-order knowledge concerning *P*, watching the second tape doesn't make her any safer with respect to *P* itself.

Perhaps the defenders of the independent check and safety arguments should pick option (2), and hold that once Jones has seen the first tape, she not only knows, but also knows that she knows that *P*. A defender of the independent check argument might hold that in any realistic case, Jones and Smith will start out with independent checks concerning the reliability of videocameras, and so Jones will be in a position to know that she knows that *P* already after watching the first tape. Along similar lines, a defender of the safety argument might hold that because videocameras are typically quite reliable, in any realistic case Jones would already be safe enough with respect

²⁸See Conee and Feldman (1998).

to P to know that she knew that P —in realistic cases, there are not only no nearby possibilities in which cameras malfunction; there are also no nearby possibilities in which there are nearby possibilities in which cameras malfunction. This way the safety theorist could allow that Jones can come to know that Smith knows that Jones knows that P without becoming safer with respect to P .

While these moves are available, they don't go to the heart of the matter. Just as it's plausible that $K_J K_S K_J P$ holds at the end of the three-stage case, it's also plausible that after a five-stage case in which Smith and Jones each watched another tape, $K_J K_S K_J K_S K_J P$ would hold, or that after a seven-stage case in which they each watched yet another tape, $K_J K_S K_J K_S K_J K_S K_J P$ would hold.²⁹ If we accept these verdicts along with closure, then we must accept that in the original case, it's not just $K_J K_J P$ that holds at the first stage, but at least $K_J K_J K_J K_J P$. The guiding idea here is that Jones' acquiring higher-order interpersonal knowledge only requires that Jones acquire more information about *Smith's* epistemic situation—not that Jones acquire more information about what she herself knows. If we endorse the independent check or safety arguments, along with closure, we'll have to reject that idea. We'll have to accept that at some point (because Jones only has finitely many independent checks, or because there are *some* possibilities in which P is false, even if they're quite far away from the possibility Jones occupies), Jones can't acquire another iteration of *interpersonal* higher-order knowledge concerning what *Smith* knows by watching a tape of Smith, because Jones doesn't have enough iterations of higher-order *intrapersonal* knowledge.

Incidentally, this is why the independent check and safety arguments imply that common knowledge is unattainable; both arguments imply that in realistic cases, while we may often know that we know, our higher-order knowledge gives out at some point. Because we always have only finitely many independent checks, or because there are always possible cases in which we are in error (even if those cases are only nearly nearly nearly nearly...nearby, for some large number of “nearly”s), we never know that we know that we know...*ad infinitum*.³⁰

I hope I've made it plausible that the intuitive verdict about our example—that watching successive tapes is a way of acquiring higher-order interpersonal knowledge—doesn't sit nicely with the independent check and safety arguments, at least if we also accept a closure principle on knowledge. In the next section I'll offer a simple formal model of knowledge that provides an attractive analysis of the case of Jones and Smith.

²⁹At least if Jones and Smith both have good enough memories to keep straight who's seen what, which we may assume they do.

³⁰For some reasons to think that this is a *bad* result—i.e., that in many cases, our mutual higher-order knowledge doesn't give out after some finite number of iterations—see Heal (1978), especially section III.

1.3 Knowledge, Information, and Normal Conditions

So far we have what looks like a puzzle. We have some *prima facie* attractive arguments against KK, along with an example in which they seem to prove too much. In the absence of some theoretical support for the example from the previous section, we might resolve this tension by simply dismissing the intuition that Jones knows that Smith knows that Jones knows that P after the third stage of this story, or at least dismissing the intuition that such a story could be extended to include arbitrarily many stages. My strategy in the remainder of this chapter will be as follows. First, I'll present a simple theory of knowledge on which KK and closure hold, and which I'll argue provides an attractive analysis of the poker example from the previous section. The fact that our intuitive reactions to the example from the previous section can be accounted for in a simple, elegant model of knowledge militates against simply dismissing the intuitions. This won't yet resolve our puzzle—we'll won't yet have an explanation of what was attractive about the arguments against KK. I'll go on to argue that when the model I'm about to discuss is combined with contextualism concerning knowledge, we can explain away apparent failures of KK.

1.3.1 Normal Conditions and KK

Here's a simple, two-prong analysis of knowledge: S knows that P just in case:

- S is in a state that carries the information that P , and
- S 's being in this state causes or constitutes S 's believing that P .

It is inspired primarily by Fred Dretske (1981), but similar views have been defended by many other writers.³¹ Moreover, as I argue below, even views that aren't all that similar to Dretske's (e.g., David Lewis' contextualist theory of knowledge) can be seen as sharing the same abstract structure, and it's this abstract structure that I'll ultimately argue leads to KK.³²

Of course, this theory is only helpful insofar as we have some idea of what it is to be in a state that carries the information that P . Here's how the phrase is intended to be understood: S is in a state X that carries the information that P just in case:

- Whenever conditions are normal, S is in X only if P , and
- Conditions are normal.

Analyzing the notion of information in terms of what would be true in certain conditions is a strategy that goes back to Dennis Stampe (1977), and which has had many proponents, including Dretske (1981), Ruth Millikan (1984, 1989), and Robert

³¹See e.g., Stalnaker (1999).

³²Of course Dretske himself was no friend of KK. But just as contextualists hold that the Dretskean rejection of closure stems from a failure to appreciate the importance of context, I'll suggest a similar diagnosis of the Dretskean rejection of KK.

Stalnaker (1999). Of course, this further analysis is informative only insofar as we have some idea of what normal conditions are. It may help to start with an intuitive example. Suppose I see a dog. Conditions are normal in that my eyes are working properly, there are no misleading stimuli in my environment (e.g., no papier-mâché dogs), etc. In such a case, my brain (in particular, my visual system) will be in some state X such that, provided conditions are normal in the ways just gestured at, it will only be in X if there is in fact a dog in front of me. On the intended interpretation of “carries the information,” in such a case my visual system carries the information that there is a dog in front of me.³³

Some authors have tried to provide reductive, naturalistic analyses of normality—Stampe (1977) endorses such ambitions, though the most developed account I know of is provided by Millikan (1984)—but we needn’t assume that a reductive analysis of normality is possible in order to be interested in the philosophical consequences of a normal-conditions theory of knowledge. For instance, we might agree with Williamson (2000) and despair of finding a non-circular analysis of knowledge, while hoping that the above analysis, though in some sense circular (because we may need to appeal to the notion of knowledge in explicating our concept of normal conditions) can still lead to some interesting conclusions about the nature of knowledge.

Moreover, it’s consistent with rejecting the hope for a *completely* reductive analysis that we try to say *some* informative things about normality. For instance, we might hold that which conditions count as normal is in some sense context-sensitive, and that our ascriptions of knowledge have to be understood as being made relative to potentially flexible background presuppositions about which conditions count as normal. Later in this chapter I’ll invoke such a hypothesis to explain away some of the apparent problems with the KK thesis.

The class of “normal conditions” theories of knowledge, as I’m understanding that class, is a large one. This is because the sense of “normal” that I’m using is quite thin. A theory counts as a normal conditions theory as long as it satisfies two conditions. First, it identifies some range of “normal” possibilities, and says that you carry the information that P just in case you’re in a state X that—as long as we restrict our attention to normal possibilities—you’re only in if P . Second, it goes on to explain knowledge in terms of information-carrying. But as I’m using the

³³This way of putting things suggests that I’m describing not normal conditions *simpliciter*, but normal conditions *for the operation of vision*. In fact, I think there are objections to the account in the text which may be best addressed by moving to a version of the normal conditions account that relativizes normal conditions to processes, such as vision. For instance, we might hold that S is in a state that carries the information that P just in case:

- S was caused to be in a state X by a process Y such that under all normal _{Y} conditions, S is in X only if P .
- Conditions are normal _{Y} .

Here, the claim that conditions are normal _{Y} is the claim that conditions are normal for the operation of process Y . The objections that would prompt moving to a process-relativized account won’t be relevant to the arguments I’ll offer in the text, so I’ll stick with the simple account for ease of presentation. While it may not be obvious, the same arguments concerning KK that I’ll present in the context of the simple account would go through on the process-relativized account as well.

phrase “normal,” a theory counts as a normal conditions theory as long as it identifies some such set—it needn’t be the set of statistically normal possibilities, or biologically normal possibilities, or possibilities that are normal in any particular pre-theoretic sense of “normal.”

In particular, pretty much any relevant alternatives theory of knowledge—that is, any theory according to which knowing that P requires being able to eliminate relevant alternatives to P , but where which alternatives count as relevant may vary from proposition to proposition, or context to context—will count as a normal conditions theory.³⁴ We can translate relevant alternatives theories into normal conditions theories by holding that whenever an alternative is “relevant,” the possibility in which that alternative is realized counts as “normal.” To see how this works, it will help to consider an example.

Here’s how David Lewis summarizes his contextualist relevant alternatives theory of knowledge:

I say S knows that P iff P holds in every possibility left uneliminated by S ’s evidence—Psst!—except for those possibilities that we are properly ignoring. (Lewis, 1999, p.436)

On Lewis’ theory, one knows that P just in case one’s evidence eliminates all of the relevant (i.e., not properly ignored) not- P alternatives. Given how Lewis understands elimination by evidence, we can translate this into normal conditions talk as follows:

S knows that P iff S is in an experiential state X such that in all normal (i.e., not properly ignored) possibilities, if S is in X then P .

The translation makes it clear that Lewis’ theory is a version of the normal conditions theory. His additions include the restriction to experiential states as the relevant type of information-carrying states, and a great deal of detail concerning which possibilities can and cannot be properly ignored (i.e., which possibilities count as normal). His theory is contextualist (like many relevant alternatives theories) because which possibilities count as relevant/normal is determined in part by context.

We could use “normal conditions” more narrowly, such that a theory only counts as a normal conditions theory if it gives some particular sort of account of normality—perhaps statistical normality, or biological normality, or something along those lines. While the distinction between normal conditions theories in the narrow sense and the broader sense might be significant for some purposes, it isn’t for mine. In the remainder of this section, I’ll argue that normal conditions theories entail KK, and my argument will apply to any normal conditions theories even when we use the broad understanding of “normal conditions” theories according to which pretty much any

³⁴The idea of a “relevant alternatives” theory of knowledge comes from Dretske (1970), but many epistemologists—especially contextualists of various stripes—have defended versions of relevant alternatives theories. Given that Dretske is the originator of both relevant alternatives theories of knowledge, and normal conditions theories, perhaps it shouldn’t be surprising that we can see one type of theory as a special case of the other.

relevant alternatives theory will count as a normal conditions theory.³⁵

To make good on this promise, however, I will need to make a simplifying assumption. In particular, in keeping with my focus on versions of the KK principle that abstract away from obstacles to KK involving failures of *belief* to iterate, I will assume that whenever agents are in states that carry the information that P , this causes or constitutes their believing that P .³⁶ Given this assumption, we can identify knowing that P with carrying the information that P , and we can ignore the belief requirement on knowledge. What, then, would it take for the KK principle to hold, on a normal conditions theory of knowledge? It would need to be the case that one's carrying the information that P entailed carrying the information that one carries the information that P .

I'll first give an informal argument to the effect that in fact this *is* the case, and then I'll provide a formal statement of the argument. Here's the basic idea. If S is in a state that carries the information that P , then *that very state* also carries the information that S is in a state that carries the information that P . Why is this? Higher-order information carrying requires that one be in a state that is correlated (given normal conditions) with being in a state that is correlated (given normal conditions) with P . But because every state is correlated with itself, if one is in a state X that is correlated with P , then one is also in a state (X itself) that is correlated with being in a state that is correlated with P .³⁷

For example, suppose the current state of my stomach carries the information that I'm hungry. My stomach is growling, and under normal conditions, my stomach only growls when I'm hungry. If my stomach is growling, then my stomach is in a state (the growling state) that carries the information that it's in a state (the very same growling state) that carries the information that I'm hungry. This is because whenever my stomach is growling, I'm in a state—the growling stomach state—that carries the information that I'm hungry. Higher-order information comes for free with first-order information carrying.

Here's the formal statement of the argument. Let N be the claim that conditions are normal, and let \Box_N be a restricted necessity operator; $\Box_N P$ means that in all N -worlds, P holds.

By our definitions, S carries the information that P just in case:

³⁵Though for reasons that will become clear, it will *not* apply to normal conditions theories that have normal conditions vary from proposition to proposition, in particular from P to $K_S P$. This is the view that Dretske himself endorsed, and it explains why he rejected KK and closure.

³⁶This assumption is obviously false; I might be in a state that carries the information that my appendix has burst (e.g., the state of having a burst appendix carries this information) without believing or knowing that my appendix has burst. But making this assumption amounts to abstracting away from the belief requirement on knowing. A defense of KK on this assumption would amount to the claim that—roughly—if the belief requirements for knowledge were always met, *then* KK would hold. Because the objections to KK I'm concerned with in this chapter are objections that turn on the idea that *even if* one has the necessary higher-order beliefs, one might still know while failing to know that one knows, this assumption is legitimate in the present context.

³⁷Thanks to Bob Stalnaker for discussion here. See also Williamson (2001), who credits Lloyd Humberstone with pointing out (in conversation) that Lewis' theory vindicates some strong principles of epistemic logic, including KK.

1. S is in a state X such that $\Box_N(S \text{ is in } X \supset P)$, and
2. N .

What does it take for an agent S to have higher-order knowledge on the present account? S must be in a state that carries the information that the above two conditions (1 and 2) are met, and conditions must be normal. That is, if S is in a state that carries the information that S is in a state that carries the information that P , then:

- 1'. S is in a state X' such that $\Box_N(S \text{ is in } X' \supset (S \text{ is in a state } X \text{ such that } \Box_N(S \text{ is in } X \supset P) \ \& \ N))$ and
- 2'. N .

To show that the present account of knowledge entails KK, it will suffice to assume that 1 and 2 hold and prove 1' and 2' from these assumptions. The proof is straightforward.

That 2' follows from 2 is obvious—they are identical. Moreover, the last “ N ” in 1' can be ignored—any claim of the form $\Box_P(Q \supset R \ \& \ P)$ is equivalent to $\Box_P(Q \supset R)$. So we are now faced with showing that 1'' (the same as 1', but with an “ N ” removed) follows from 1:

1. S is in a state X such that $\Box_N(S \text{ is in } X \supset P)$
- 1''. S in a state X' such that $\Box_N(S \text{ is in } X' \supset S \text{ is in a state } X \text{ such that } \Box_N(S \text{ is in } X \supset P))$

If we let $X' = X$, we can see that 1'' follows from 1. We can rewrite it a third time, substituting X for X' in 1'':

- 1'''. S in a state X such that $\Box_N(S \text{ is in } X \supset S \text{ is in a state } X \text{ such that } \Box_N(S \text{ is in } X \supset P))$

1''', while hard to parse, is equivalent to 1. Our proof is now complete—on the present account, carrying the information that P (knowing that P) entails carrying the information that one carries the information that P (knowing that one knows that P).

The reasons why closure holds, in this model, are very similar to the reasons why KK does. Suppose P entails Q . Then if S is in a state that carries the information that P , S is in a state that carries the information that Q —if under normal conditions S is in X only if P , and P entails Q , then under normal conditions S is in X only if Q . In this model, knowing that P entails knowing that Q , if Q follows from P .

The example of the growling stomach may help again. Suppose that it follows from the fact that I'm hungry that I'm less than fully comfortable. If my growling stomach carries the information that I'm hungry—i.e., if under normal conditions, my stomach only growls when I'm hungry—then it must carry the information that

I'm less than fully comfortable. If under normal conditions, my stomach only growls when I'm hungry, then under normal conditions, my stomach only growls when I'm less than fully comfortable.

This model has some counterintuitive features—KK and closure themselves can seem counterintuitive, and the model also implies that agents trivially know that conditions are normal whenever conditions are in fact normal.³⁸ Some counterintuitive consequences of the model can be deflected by pointing out that the simplifying assumptions I made is false; often when one carries the information that *P*, one nevertheless doesn't *believe* that *P*, and so won't count as knowing that *P* according to the model I've been discussing. In the final section of the chapter I'll introduce another, general strategy for blunting counterintuitive consequences of the model—as a preview, we can hold that which conditions count as normal needn't be the same from context to context, and that apparent failures of principles like KK and closure can be explained as cases in which focusing on certain questions induces a shift in context. Before introducing these complications, however, it will be fruitful to see how we might apply the current model—KK, closure, and all—to the example from the previous section.

In the poker example, it's natural to hold fixed what we mean by normal conditions; conditions are normal just in case the camera is functioning properly, and it's commonly believed (or merely commonly implicitly trusted) that the camera is functioning properly.³⁹ As argued above, the present account of knowledge entails KK and closure when normal conditions are held fixed. So the present account can explain why our initial description of the three-stage case was correct. Upon seeing the original tape, Jones comes to believe that she has the stronger hand, and because conditions are normal, and normally she'll only have this camera-based belief if she *really does* have the stronger hand, she knows that she has the stronger hand. At the second stage, Smith acquires a camera-based belief that Jones knows that she has the stronger hand. Again, if conditions are normal—the camera is functioning properly, and this is commonly believed—then Smith will only have this higher-order camera-based belief if Jones really does know that she has the stronger hand. And so on at the third stage, and potentially beyond—the present account has it that seeing still further tapes would supply ever more iterations of interpersonal knowledge.

Because the account entails KK, it avoids positing obstacles to Jones' coming to know that Smith knows that Jones knows that *P* associated with obstacles to Jones' coming to know that she herself knows that *P*—because KK holds, knowing that one knows doesn't require some independent check, or extra safety with respect to *P*. Because (intuitively) Jones *can* come to know that Smith knows that Jones knows that *P* just by watching another tape (and without having an independent check or getting any more safe with respect to *P* itself), this is an attractive feature of the present account.

³⁸Any state you could be in is a state such that, if conditions are normal and you're in that state, then conditions are normal.

³⁹The definition of "common belief" is what you get when you substitute "belief" for "knowledge" in the definition of common knowledge mentioned at the beginning of this chapter.

We still, however, haven't resolved our puzzle. Why is it that in some cases—like the case of Jones and Smith—it seems natural to think that higher-order knowledge doesn't require independent checks or extra safety, while in other cases—e.g., the case of the reliable chicken-sexer—it seems natural to think that it does? In the next section I'll offer a contextualist attempt at a reconciliation, along the lines of standard contextualist strategies for saving closure from apparent counterexamples.

1.4 KK, Closure, and Context

It's a standard part of contextualist lore that while closure holds *within* a context, it does not hold *across* contexts.⁴⁰ Hold fixed the context, and if I know that P , then I know that Q , if P entails Q . But it may be that once the question of whether Q is raised, I no longer speak truly by self-ascribing "knowledge" of Q , since raising the question shifts the context, and thereby shifts the extension of "knowledge."⁴¹

For instance, consider Dretske's (1970) example of a man and his son looking at a zebra at the zoo. Dretske rejected closure, and held that in such a situation, one might know that the animal in the enclosure is a zebra, while at the same time failing to know that it is not a cleverly disguised mule. One motivation for this position comes from reflection on how it would be reasonable for the man to respond to various questions. For instance, in response to his son's question: "Do you know what kind of animal that is?" it would be quite reasonable for the man to respond: "Why yes, it's a zebra." But in response to the question: "How do you know that it is not a cleverly disguised mule?" it would be reasonable to respond: "I know no such thing; I am in no position to rule out that possibility." If we agree with these responses and take them at face value, as Dretske does, then it seems as if we must reject closure. After all, an animal's being a zebra entails its not being a cleverly disguised mule, so closure would imply that if one knows that an animal is a zebra, then one knows that it is not a cleverly disguised mule.

Contextualists have a different account of what's going on in cases like the one Dretske discusses. According to contextualists, in response to the initial question as to what kind of animal is in the enclosure, the man can speak truly by claiming to "know" that it is a zebra. In this context, before the question about cleverly disguised mules has been asked, the man "knows" not only that the animal is a zebra, but also that it is not a cleverly disguised mule; closure holds. But once the question concerning cleverly disguised mules has been asked, the extension of "knowledge" shifts, and the man speaks truly by denying that he "knows" that the animal is not a cleverly disguised mule. He would also speak truly by denying that he "knows" that the animal is a zebra—again, closure holds in the new context as well. Closure won't hold across contexts, however; what follows from what's "known" in the first context—e.g., that the animal in the enclosure isn't a cleverly disguised

⁴⁰Stine (1971) defends an early version of this strategy as a response to Dretske (1970). For some other examples, see Cohen (1987) and Lewis (1999).

⁴¹From here on out I'll be sloppy about use and mention, and to avoid distracting the reader I will omit quotes in cases where, strictly speaking, they would be appropriate.

mule—needn't be "known" in the second.

If we understand contexts as determining sets of normal conditions, then we have a nice explanation of this—as argued above, when normal conditions are held fixed, knowledge is closed under logical implication. When normal conditions vary, however, it needn't be. P might hold in all normal conditions N , and P might entail Q , but raising the question of whether Q might bring us to a context with a new set of normal conditions N' , in some of which neither P nor Q holds.

Why opt for the contextualist account of the case rather than the simple one on which closure really does fail? One reason comes from the phenomenon of "abominable conjunctions" (DeRose, 1995). While it can seem reasonable to say that the man knows he's looking at a zebra, and then to say that he doesn't know that he's not looking at a cleverly disguised mule, there seems to be something wrong with saying both in one breath. The following conjunction is abominable:

(AC) The man doesn't know that the animal is not a non-zebra cleverly disguised to look like a zebra, but he does know that it's a zebra.

If we deny closure, it's hard to explain why anything should be wrong with sentences like (AC). But if we are contextualists, we can hold that while there are contexts in which "he knows it's a zebra" is true, and contexts in which "he doesn't know that it's not a non-zebra cleverly disguised to look like a zebra" is true, there is no context in which their conjunction, (AC), is true.

While diagnosing putative failures of closure along the above lines is by no means uncontroversial, strategies like it have seemed attractive to many contextualists. We might try to use a similar strategy, however, to diagnose putative failures of KK. Such a strategy would involve holding that within a context (i.e., holding fixed normal conditions), KK holds. But in a context in which S knows that P , shifting our focus to the question "Does S know that S knows that P ?" might bring us into a new context in which S has neither second nor first-order knowledge.⁴²

There are at least two questions we should ask before endorsing such a strategy. First, what is the *prima facie* motivation for pursuing it? Second, how will the details work—that is, why should asking questions about higher-order knowledge shift the conversational context in ways so as to make KK seem false?

⁴²In recent years there have arisen rivals to contextualism that seek to explain much of the same data that contextualism is meant to explain. I have in mind subject sensitive invariantism of the sort defended by John Hawthorne (2004) and Jason Stanley (2005), along with relativism of the sort defended by John Macfarlane (Forthcoming). It's not clear to me whether the issues that divide these newer theories from contextualist theories of the traditional sort are ones on which I need to take a stand for my present purposes. While I've focused on contextualist theories here, it may be that many or all of the same points could be made in a sensitive invariantist or relativist framework. In particular, I suspect that each of the places I appeal to context relativity are places that a relativist could appeal to as assessor or standard relativity.

1.4.1 Defending the Parallel

The *prima facie* motivations for pursuing a contextualist defense of KK are very similar to those that militate in favor of a contextualist defense of closure. Firstly, there are theoretical motivations. As I've already argued, normal conditions theories of knowledge entail closure and KK alike, when normal conditions are held fixed. If we defend closure by holding that apparent counterexamples involve context shifts, we should take a parallel line concerning KK.

But we can also appeal to the phenomenon of abominable conjunctions to motivate KK. Return to the case of the reliable chicken-sexer, who we may name "Jimmy." Suppose Jimmy in fact has a reliable chicken-sexing faculty, and this faculty produces in him the belief that some chicken is female. Should we say that Jimmy knows that the chicken is female? In some contexts, this may seem perfectly reasonable. However, suppose we go on to urge that Jimmy doesn't know that he knows the sex of the chicken, because he can't rule out the possibility that he has no reliable chicken-sexing faculty and is merely guessing. Once this is done, it's a bit awkward to insist that he nevertheless does know that the chicken is female—if Jimmy can't rule out the possibility that he's just guessing, then how can he rule out the possibility that the chicken isn't the sex he thinks it is? That is, how can he know what sex the chicken is? One way of bringing out the tension involves considering the following abominable conjunction:

(AC') Jimmy doesn't know whether he knows the sex of the chicken or is merely guessing, but he knows that he's not guessing incorrectly.

The opponent of KK must accept the first conjunct because she denies that Jimmy knows that he knows. Why must she grant the second conjunct? Suppose Jimmy knows that the chicken is female. It follows from the fact that the chicken is female, and from the fact that Jimmy believes that the chicken is female, that Jimmy isn't guessing incorrectly. So if the opponent of KK grants that Jimmy knows that the chicken is female, and grants that Jimmy knows that he believes that the chicken is female, and grants closure, then she must grant that Jimmy knows that he's not guessing incorrectly.⁴³

If we deny KK, then we'll accept that one sometimes knows that *P* without knowing whether one knows or is merely guessing. But then we can't explain why conjunctions like (AC') are abominable—we can't explain why there's a tension between acknowledging that for all Jimmy knows, he might be guessing (i.e., he doesn't know that he knows the sex of the chicken), *and* insisting that he nevertheless does

⁴³If the opponent of KK rejects closure, she can also resist the conclusion that that Jimmy knows that he's not guessing incorrectly. The argument that Jimmy knows that he's not guessing incorrectly implicitly relied on closure. If closure fails, then Jimmy might know that he believes that the chicken is male, and know that the chicken *is* male, but fail to know whether his belief about the maleness of the chicken is correct. I take it, however, that appealing to closure in the present context is dialectically appropriate. After all, I'm arguing that there are underappreciated similarities between closure and KK, and that arguments for closure are mirrored by similar arguments for KK. My targets, then, are those who are already sympathetic to closure.

know the sex of the chicken. The contextualist theory, on the other hand, provides a satisfying explanation of how such a tension arises; there's no one assignment of normal conditions that makes true both the claim that Jimmy knows the sex of the chicken, *and* the claim that he doesn't know that he knows. We can still explain why this doesn't seem obvious by holding that shifting our focus between the two questions tends to shift the sets of conditions that count as normal for the purpose of knowledge ascriptions. My strategy has been exactly analogous to arguments that have been offered on behalf of closure; just as denying closure leads to abominable conjunctions, so does denying KK.

Another parallel between extant contextualist defenses of closure and potential contextualist defenses of KK concerns the role of closure and KK in skeptical arguments. A common skeptical strategy is to appeal to some putative piece of knowledge that we lack—the knowledge that we are not being deceived by an evil demon, perhaps—and to then appeal to a closure principle to conclude that we lack other pieces of knowledge that we ordinarily take ourselves to have. Some antiskeptics will respond to this style of argument by rejecting closure. Others will respond by insisting that we *do* know that we're not being deceived by an evil demon. But contextualists will offer a diagnosis much like the one I've already discussed concerning the case of the cleverly disguised mule—we can accept both skeptical claims to the effect that we lack certain pieces of knowledge, and ordinary commonsense claims to the effect that we have certain pieces of knowledge, *and* closure. According to the contextualist, the apparent conflict between commonsense knowledge attributions and skeptical denials of knowledge in fact stems from a failure to appreciate the fact that such attributions and denials are typically made in different contexts, in which different standards for knowledge are at play.

There is a parallel dialectic concerning KK and skepticism. Certain skeptical arguments start with the denial that we have some piece of higher order knowledge. The skeptic might argue, for instance, that we don't know that we know that we have hands, perhaps on the grounds that such higher-order knowledge would require being able to meet certain challenges that we are in fact unable to meet. She can go on to argue that, because KK holds, if we lack higher-order knowledge, we also lack first-order knowledge—we don't know that we have hands. Adler (1981) and Stroud (1984) both consider arguments in this family, Stroud endorsing them, and Adler rejecting them on the grounds of opposition to KK. Contextualism suggest a natural strategy for avoiding full-blown skepticism, while still holding onto KK. Claims to the effect that we have first-order knowledge, and other claims to the effect that we *lack* higher-order knowledge, can both be true even if KK holds, so long as such claims are made in different contexts. The contextualist treatment of skeptical arguments that rely on KK, then, exactly parallels the contextualist treatment of skeptical arguments that rely on closure.

1.4.2 Context and Higher-Order Knowledge

Now that I've shown that the *prima facie* motivations for pursuing a contextualist defense of KK are exactly analogous to those that motivate pursuing a contextualist

defense of closure, we can turn to the second question: why should asking questions about higher-order knowledge shift the context in ways so as to make KK seem false?

One simple explanation of why this might happen would appeal to salience. That is, we might hold that if some possibility is salient, then it should be included in normal conditions, and also hold that focusing on questions about higher-order intrapersonal knowledge typically makes new error possibilities salient.⁴⁴ While I think this is right as far as it goes, there's more that we can say.

If KK holds in all contexts, then asking whether somebody knows that she knows (after it's granted by all conversational participants that she knows) amounts to asking a question whose answer already follows from what's commonly accepted by the parties to the conversation. Since we generally try to interpret people so that the answers to their questions don't straightforwardly follow from what's already commonly accepted—plausibly, this is a special case of “accommodation” in the sense of Lewis (1979)—we may interpret the question as implicitly inviting us to shift to a new context in which more possibilities count as normal.

Why would this make for a natural interpretation of the question? Even if we grant that somebody is in a state that's correlated with P in all “normal” possibilities N , it will typically be a non-trivial question whether she is in a state that is correlated with P in all possibilities N' , where N' is a superset of N . So if we think that questions are typically interpreted so that their answers aren't trivial consequences of what conversational participants already accept, and we're sympathetic to normal-conditions accounts of knowledge, we should think that the question of whether somebody knows that she knows (asked in a context in which it's granted that she knows) will typically induce a shift into a new context where more possibilities count as normal, since this will be required if the question is to be interpreted as non-trivial.

For instance, in the case of the reliable chicken-sexer, when focusing on the first-order question of whether the chicken-sexer knows the sex of the chicken, we may presuppose a set of normal conditions that only includes cases in which her chicken-sexing faculty is reliable, and we'll therefore grant that she knows. But suppose we now raise the question of whether she knows that she knows, and we try to suggest that it should get a negative answer—after all, by her lights, she might just be guessing! Such a shift in focus may bring us to a new context, where some conditions in which she does *not* have a reliable chicken-sexing faculty count as normal. It won't follow from the fact that the chicken-sexer counted as knowing the sex of the chicken in the old context that she will count as knowing that she knows in the new context—this is so even though, in the new context, there's no difference between knowing and knowing that one knows. For the chicken-sexer to be right in all the cases that count as normal in this new context (and thereby know, and know that she knows) might require, e.g., having an independent check on the reliability of her chicken-sexing faculty. This would explain the appeal of the independent check argument—raising questions about higher-order intrapersonal knowledge may often bring us into a context in which higher-intrapersonal knowledge (and first-order

⁴⁴See Lewis (1999).

knowledge) requires independent checks.

Once we supplement our normal conditions theory of knowledge with contextualist machinery, we can also explain away the appeal of the safety argument. We must deny that safety is a requirement on knowledge if we endorse a normal conditions account of knowledge; the safety requirement is inconsistent with KK, and normal conditions accounts entail KK (holding context fixed). But we can explain why safety seems like a plausible requirement in the following manner. Suppose we try to find an intuitively plausible counterexample to safety. This would require identifying a possibility w , a subject S , and a proposition P such that it seems to us as if S knows that P in w , even though there is a nearby and relevantly similar possibility w' which S falsely believes that P .⁴⁵

But if explicitly considering a possibility tends to make that possibility salient, and thus induces a context shift so that w' counts as a “normal” possibility, then we shouldn’t expect to be able to find such a counterexample. Once we explicitly consider the possibility w' , that possibility will count as normal. Because, in w , S is not in a state that allows her to rule out the relevantly similar possibility that she is in w' , in the new context in which w' counts as a normal possibility, S will not count as knowing that P in w .

Counterexamples to the safety requirement, then, might be “elusive,” in the sense described by Lewis (1999). While in any given context, there may be instances of the safety requirement that are false, explicitly considering those instances will typically bring us into a new context in which they are true. So it is unsurprising if the safety requirement seems true—its instances will typically be such that, once they are explicitly considered, they *are* true. Incidentally, if this strategy for explaining away the appeal of the safety requirement is successful, it has implications beyond the debate over KK. The safety requirement is a crucial premise in Williamson’s enormously influential “anti-luminosity” argument, which he and others have put to use in a wide variety of philosophical contexts.⁴⁶

An additional reason for favoring a contextualist defense of KK against the independent check and safety arguments is that it has the attractive feature of explaining why higher-order *interpersonal* knowledge typically *doesn’t* strike us as requiring independent checks on the reliability of our first-order belief-forming faculties, or extra safety concerning first-order propositions. Even if it’s granted by all conversational participants that S knows that P , in a normal conditions account of knowledge, nothing follows concerning what S knows about what *other* people know. So in a context in which it’s presupposed that Jones knows that P , asking whether Jones

⁴⁵In describing what would be required for a counterexample to safety here, I am a bit more faithful to Williamson’s version of safety than I was earlier in the chapter—I refer to “relevantly similar” nearby possibilities, and require that they not just be possibilities in which P is false, but in which the subject falsely believes that P . When we do not make these qualifications, finding counterexamples to safety is a good deal easier. See Neta and Rohrbaugh (2004). Even making these qualifications, there might be other plausible counterexamples to safety. See e.g., Weatherson (2004), and Berker (2008). The argument in the text can be seen as a strategy of last resort. While we might be able to find plausible counterexamples to safety, even if we can’t, we have good reason to think that counterexamples, if they existed, might be elusive.

⁴⁶See Williamson (2000, chapter 4).

knows that Smith knows that P will typically *not* shift the context—the answer to this question needn't trivially follow from what is already presupposed, so there is no pressure to shift to a new context in response to the question. If the appeal of the idea that higher-order knowledge requires independent checks or extra safety comes from context shifts, then we can explain why those ideas are not appealing in cases in which we are discussing *interpersonal* higher-order knowledge; it's because there's no pressure to shift the context when we ask questions about higher-order interpersonal knowledge.

Incidentally, my discussion thus far has been quite concessive to the opponent of KK, in the following respect. As I mentioned at the end of §2.2, in most real life examples, it's *not* intuitive that higher-order intrapersonal knowledge requires independent checks or extra safety, because it's not at all clear how to make sense of questions about whether people know that they know, given that they know. Cases like that of the unwitting but reliable chicken-sexer are unusual precisely because there is a natural “fallback” set of normal conditions that we can move to when questions about higher-order knowledge are raised. My contextualist explanation of the “apparent” failure of KK applies only to such cases—cases in which it's clear how to move into a new context that will make questions about higher-order knowledge legitimate, even after having granted (in the old context) claims about first-order knowledge. In cases where there is no natural set of fallback normal conditions—most cases, I suspect—the contextualist strategy I've been exploring suggests that it should be hard to interpret questions about whether people know that they know; it should be hard to hear such questions as legitimate requests for information.

This prediction is born out when we consider questions not about second-order intrapersonal knowledge, but about intrapersonal knowledge of even higher orders. As I've already suggested, it is not always easy to interpret questions like the following: “I grant that Jane knows what time the movie starts, but does she know that she knows what time the movie starts?” But while such questions are at least sometimes intelligible, it is much rarer that we can make sense of analogous questions concerning third or fourth-order knowledge, let alone knowledge of even higher orders. Without bringing in heavy duty philosophical theory, there's no natural way to interpret questions like the following: “I grant that Jane knows that she knows that she knows what time the movie starts, but does she know that she knows that she knows that she knows what time the movie starts?” This should be puzzling to advocates of the independent check or safety arguments; on those views, the distinction between first-order and second-order knowledge is of a kind with the distinction between second-order and third-order knowledge, or between third-order and fourth-order knowledge; in each case, achieving the next iteration of knowledge just requires more of the same (either more independent checks, or more safety). Such views have a tricky time explaining why it should be any harder to interpret questions about knowledge of very high-orders than it is to interpret questions about second-order knowledge.

The arguments of the present chapter, however, suggest a natural explanation. In many situations, there are at least two different natural choices for how strict a set of standards we might hold knowledge attributions to (in the present framework,

a stricter set of standards corresponds to a wider set of normal possibilities). We can exploit this fact to interpret questions that presuppose the failure of KK; we can make sense of such questions by associating higher-order knowledge attributions with stricter standards than first-order knowledge attributions. But there needn't be an infinite hierarchy of ever stricter standards that could provide natural interpretations for questions about intrapersonal knowledge of arbitrarily high orders. If it were in the nature of knowledge itself that each iteration of higher-order knowledge were more difficult to acquire than the previous one, it's hard to see why there shouldn't always be available such a hierarchy to provide interpretations for such questions. But if apparent failures of KK stem from context shifts in the manner that I've suggested, then it's not so surprising that our ability to interpret questions about higher-order intrapersonal knowledge gives out relatively quickly—at least when we try to presuppose that higher-order knowledge doesn't come for free. The reason is that in most situations, there are only finitely many natural choices for how strict a set of standards to apply to knowledge attributions—only finitely many “fallback contexts” to which we might move when forced to consider a question about knowledge of a still higher order.

1.5 Conclusions

In this chapter I've offered both defensive and offensive strategies to KK sympathizers. When playing defense, a KK sympathizer can exploit the analogy to closure, and avoid counterexamples by adopting the position that KK holds within contexts, but can fail across contexts. When playing offense, a KK sympathizer can argue that her theory allows her to explain various phenomena that are puzzling when we deny KK. First, she can account for the example from §2.3—she can explain why higher-order interpersonal knowledge doesn't seem to require independent checks or extra safety, even though higher-order intrapersonal knowledge *seems* (at least sometimes) to require meeting such further conditions. Second, she can straightforwardly explain the abominability of conjunctions like (AC') by holding that such conjunctions are false in every context.

We can distinguish between easier-to-achieve and harder-to-achieve epistemic statuses. In the contextualist framework, this distinction is reflected in the fact that it's harder for a belief to count as knowledge in some contexts than it is in others. We can also distinguish between first-order knowledge and higher-order knowledge. Many contemporary epistemologists run these two distinctions together, holding that the latter distinction is a special case of the former; according to them, first-order knowledge is comparatively easy to get, second-order knowledge harder, third-order knowledge still harder, and so on all the way up. Sometimes—e.g., when discussing unwitting but reliable chicken sexers—running these distinctions together is very natural, and largely harmless. But other times—e.g., when discussing higher-order interpersonal knowledge, and especially common knowledge—conflating these distinctions gets in the way of fruitful theoretical projects; if we hold that higher-order intrapersonal knowledge is harder to achieve than lower-order intrapersonal knowledge, we

erect barriers to higher-order *interpersonal* knowledge that we may not have intended to put in place. Once we see that one can recognize both distinctions without conflating them—the contextualist approach I’ve defended does just that—we should take seriously the possibility that KK might be OK.

Chapter 2

A Puzzle About Epistemic Akrasia

2.1 The Puzzle

2.1.1 Thesis

Epistemic akrasia is irrational. By epistemic akrasia I mean the state somebody is in when she believes something of the following form:

- P , but I oughtn't believe that P .

Not much turns on the particular wording above.¹ The following formulations would do just as well:

- P , but my total evidence does not support the belief that P .

Or:

- P , but it's unreasonable/irrational for me to believe that P .²

Here's a case that will give a better idea of the sort of state I have in mind:

MATT: Matt is extremely afraid of flying. When professional obligations require him to travel (even thousands of miles), he either drives or takes a train. He does not travel overseas. When his friends and loved ones travel by air, he obsessively checks the status of their flights online, and calls them as soon as possible after landing to make sure that they're OK. When asked about all this behavior, he doesn't defend it. Instead, he says things like the following: "Of course the evidence shows that flying is not

¹I'll alternate freely between the locutions "one ought to believe that P ," "it is reasonable to believe that P ," "one has justification to believe that P ," "it is rational for one to believe that P ," and "one's evidence supports believing that P ." I don't deny that these locutions can fruitfully be given distinct senses, but I do think that this is a matter of stipulation; I don't think ordinary language already distinguishes them.

²These claims are all reminiscent of Moore's paradox. I believe there are close connections between Moore's paradox and the phenomenon of epistemic akrasia, but discussing Moore's paradox in any depth is beyond the scope of this dissertation.

particularly dangerous—certainly less dangerous than driving comparable distances, but I just can’t shake the belief that if I fly, my plane will crash and I will die. What’s holding it up there anyway?”

Though not completely uncontroversial, it’s natural to describe Matt’s psychological state as one in which he believes that flying is dangerous, but at the same times believes that this belief is irrational.³

I call this state “epistemic akrasia” because of the close structural similarity it bears to states of akrasia more familiar from the practical realm, and in the hopes that drawing attention to this similarity will reinforce the idea that epistemic akrasia is irrational.⁴ Just as there seems to be something practically irrational about performing an action while believing that one oughtn’t perform that action (e.g., eating the donuts while believing that one should stick to one’s diet), there seems to be something epistemically irrational about believing some claim while believing that one’s total evidence does not support that claim (e.g., believing that one’s plane will crash, while believing that one’s belief is the product of an irrational phobia).⁵

In fact, practical akrasia can be a symptom of epistemic akrasia. Matt will presumably turn down opportunities to travel by plane, (he believes he’ll die if he flies) while at the same time cursing himself for making such foolish, irrational decisions (he believes that the rational decision in his situation is to travel by plane). Once you’re epistemically akratic, it’s often hard to avoid being practically akratic as well.

Many philosophers simply take for granted that epistemic akrasia is irrational, or at least take the claim that it is irrational to be in need of only minimal defense. T.M. Scanlon takes epistemic akrasia to be a paradigm case of irrationality:

Irrationality in the clearest sense occurs when a person’s attitudes fail to conform to his or her own judgments: when, for example, a person continues to believe something (continues to regard it with conviction and to take it as a premise in subsequent reasoning) even though he or

³Some philosophers have argued that epistemic akrasia is impossible (see Hurley (1989), Pettit and Smith (1996), Adler (1999, 2002), and Raz (2007)), so there are no cases that can be completely uncontroversially described as ones in which someone holds a belief while taking that belief to be irrational. For instance, rather than describing this case in which Matt straightforwardly believes that flying is dangerous, and straightforwardly believes that he oughtn’t believe this, we might instead describe it as a case of “in-between believing”; one in which Matt neither straightforwardly believes nor straightforwardly disbelieves that flying is dangerous. See Schwitzgebel (2001, 2002). While this isn’t how I put things, I suspect that such a position would ultimately dovetail nicely with the theses defended in this chapter.

⁴Though I’m certainly not the first to discuss such states, or to use the phrase “epistemic akrasia.” See, e.g., Owens (2002).

⁵Of course, depending on the specific nature of the case, it might be reasonable to revise one’s beliefs in either direction—paradigm cases are ones in which the first-order belief or intention ought to be revised, but there will be others in which the first-order belief or intention is fine, and it is the normative judgment about that belief/intention that should be revised. For instance, we might say that Huck Finn is akratic with respect to the question of whether he should turn Jim in, but that he ought to revise his belief about the moral status of slavery, rather than changing his intention not to turn Jim in. See (Arpaly, 2003).

she judges there to be good reason for rejecting it...(1998, p.25)⁶

However, as the following section will bring out, there are some general theoretical reasons for thinking that epistemic akrasia must sometimes be rational.

2.1.2 Antithesis

Justified false beliefs are possible. This claim commands *almost* universal agreement among contemporary epistemologists.⁷ Most would go further, and claim that the possibility of justified false belief isn't restricted to special subject matters. Rather, in no domain is justification an infallible guide to truth.⁸ At the very least, if there *are* domains in which justification entails truth, this requires some special explanation.

Here's an argument to the conclusion that in no domain is justification an infallible guide to truth. Testimony seems to be capable of generating justified beliefs in any domain. One can come to hold justified mathematical beliefs by relying on testimony from expert mathematicians, one can come to hold justified gastronomical beliefs by relying on testimony from master chefs, and so on. But testimony is a fallible process—a generally reliable testifier might on some particular occasion make a mistake, or decide to tell a lie. In such cases, the hearer of testimony will end up with a justified, false belief. This process can work in any domain; even in domains where our knowledge might seem most secure, it can plausibly be undermined by misleading testimony.⁹

However, if we hold that epistemic akrasia is always irrational, we must reject (or at least qualify) these claims. For the claim that epistemic akrasia is always irrational is really just the claim that a particular sort of justified false belief—a justified false belief concerning what one ought to believe—is impossible. To make this clear, it will help to consider a concrete example.¹⁰

Imagine an undergraduate who takes an epistemology course in which she is exposed to apparently flawless skeptical arguments. While the arguments must contain flaws—I assume here that skepticism is false—her professor presents them as conclusive, and the student is in no position to see what's wrong with them. Suppose the student, on the basis of her exposure to the skeptical arguments, comes to believe that she oughtn't believe that she has hands. This belief is false—the skeptic is wrong—but it's tempting to think that the student is justified in believing this

⁶See also Feldman (2005) and Smithies (Forthcoming), both of whom take the claim that epistemic akrasia is irrational to be relatively obvious.

⁷Though see Clayton Littlejohn's "The Myth of the False, Justified Belief" (Ms) for a dissenting note.

⁸This isn't to say that there are no *propositions* such that justifiably believing those propositions entails truly believing them. Necessarily true propositions would provide counterexamples to that claim; if one justifiably believes that $2 + 2 = 4$, then one truly believes that $2 + 2 = 4$. Rather, the claim relies on a the vague (but intuitive) idea that propositions can be grouped into *domains*, and that for no domain D is it the case that for all propositions P in domain D , justifiably believing P entails that P is true.

⁹Field (2005) and Christensen (2007) both discuss the possibility of mathematical knowledge being undermined by misleading testimony.

¹⁰This example is drawn from Feldman (2005).

falsehood, given that she's been exposed to powerful (albeit misleading) arguments in favor of it. Let's also suppose that the student's belief that she has hands remains unshaken. That is, the student believes both halves of the following conjunction:

- I have hands, but I oughtn't believe that I have hands.

The student, then, is epistemically akratic. But if it's possible to have justified false beliefs concerning whether or not you ought to believe you have hands, then it's hard to see how the student could be irrational for believing as she does. After all, she's responding appropriately to her evidence in believing she has hands, (skepticism is false, so that's what she should believe) *and* she's responding appropriately to her evidence in believing that she oughtn't believe that she has hands (because she has misleading evidence concerning the truth of skepticism).¹¹

Here's another way of putting the point. The proposition that you have hands and the proposition that you ought to believe that you have hands seem to concern different subject matters; one concerns anatomy, and the other concerns epistemology. And it's hard to see why one's evidence about anatomy and one's evidence about epistemology must always march in lock step—couldn't one have evidence for an anatomical claim (that you have hands) while lacking evidence for an epistemological one (that you ought to believe that you have hands)?

In the face of considerations like these, we are faced with two options:

1. Concede that epistemic akrasia is sometimes rational.
2. Maintain that epistemic akrasia is always irrational, and try to explain why beliefs concerning what one ought to believe should be a counterexample to an otherwise compelling generalization concerning the possibility of justified false belief.

A number of philosophers have recently endorsed versions of option 1.¹² If we endorse option 1, we take on the burden of explaining away the temptation to think that epistemic akrasia *must* be irrational—that it is a species of inconsistency or incoherence. In my opinion, this burden cannot be discharged—I'll ultimately argue that epistemic akrasia is actually quite closely related to straightforwardly inconsistent belief.

¹¹Weatherson (2008) offers a similar argument, though he focuses on the general possibility of justified false beliefs about epistemology, and less on the specific case of testimony as a method for generating such beliefs.

¹²See Weatherson (2008) and Wedgwood (2011), and arguably Williamson (2000) and Lasonen Aarnio (Forthcoming), for endorsements of the view that epistemic akrasia is sometimes rational. Williamson and Lasonen Aarnio are tricky cases, since both are sympathetic to the idea that one ought to believe that *P* only if one knows that *P*. Given this view, it can never be the case that one ought to believe that one ought to believe that *P*, while one oughtn't believe that *P*—knowledge is factive, so if one knows that one knows that *P*, then one knows that *P*. But because they allow for cases in which one ought to be confident that one ought to be confident that *P* even though one oughtn't be confident that *P*, I think they deserve to be included along with Weatherson and Wedgwood as defenders of the rationality of (something like) epistemic akrasia.

Other philosophers have continued to maintain that epistemic akrasia is irrational in the face of examples like the one above, but (as the wording of option 2 makes clear) they bear an explanatory burden as well. It's not enough to simply point out that epistemically akratic agents will say odd things and act in odd ways—if we hold that epistemic akrasia is always irrational, we're owed a story about what's so special about epistemological beliefs that they should be a counterexample to otherwise compelling general principles. In the remainder of this chapter, I'll try to provide such a story.

2.2 What is Epistemic Akrasia?

I believe that in order to make progress on the question of whether epistemic akrasia might be *rational*, we need to get clearer on the question of what epistemic akrasia *is*. In particular, what does somebody have to be like in order to be legitimately describable as both believing some claim, and believing that she oughtn't believe that claim? While examples like MATT give us some guidance, can we say anything more general about what epistemic akrasia involves?

Because akrasia has received more discussion in the practical realm than the epistemic one, it will be helpful to try to piggyback on accounts of akrasia that focus on the practical case. But first a warning: the accounts of akrasia which I'll discuss are not completely neutral with respect to metaethical questions about the nature of moral judgment. Because of this, building an account of epistemic akrasia by analogy to such accounts will not be completely neutral with respect to metaepistemological questions about the nature of epistemological judgment.

A popular strategy for making sense of akrasia in the practical case involves understanding it as a species of inner conflict, or fragmentation. If we take this strategy, we'll hold that an agent who is akratic with respect to some action ϕ has conflicting dispositions vis-à-vis ϕ , perhaps produced and/or maintained by distinct inner processes or selves.¹³

One way of developing this theme involves emphasizing the role of moral emotions.¹⁴ We might hold that in paradigm cases, when an agent judges that she ought to ϕ , she is both disposed to ϕ , and disposed to feel various moral emotions; perhaps she is disposed to feel guilty upon failing to ϕ , and to blame others for failing to ϕ . However, it is possible, so the story will continue, for an agent to be disposed to feel these emotions *without* being most motivated to ϕ . We might imagine that

¹³John Stuart Mill was an early defender of such a strategy. He describes a case of akrasia as follows:

But it is obvious that "I" am both parties in the contest; the conflict is between me and myself; between (for instance) me desiring a pleasure, and me dreading self-reproach. What causes Me, or, if you please, my Will, to be identified with one side rather than the other, is that one of the Me's represents a more permanent state of my feelings than the other does. After the temptation has been yielded to, the desiring "I" will come to an end, but the conscience-stricken "I" may endure to the end of life. (Quoted in Dennett 1984, pp.90-91.)

¹⁴See esp. Blackburn (1998, §3.3).

even though her emotions have *some* motivational force, they are outweighed by a powerful urge not to ϕ , and so that all things considered, when given the opportunity, the agent will fail to ϕ . Moral akrasia, then, would be understood as involving a divergence between what one's moral emotions motivate one to do, and what some other psychological system (whichever produces urges) motivates one to do.

This strategy doesn't generalize nicely to provide an account of epistemic akrasia. In the moral case, we have a wide array of emotions to appeal to—negative ones like anger, resentment, guilt, and contempt, along with positive ones like admiration and gratitude.¹⁵ But in the epistemological case, we don't seem to have a set of emotions that plausibly have constitutive links with epistemological judgment. To provide an account of epistemic akrasia parallel to the account of moral akrasia just sketched, we would have to first identify a set of distinctively epistemic emotions. Having done this, we could then hold that epistemic akrasia involves cases in which an agent's epistemic emotions dispose her to believe one thing, while some other psychological subsystem disposes her to believe another. In the absence of a story about epistemic emotions, however, this strategy doesn't look promising.

This points to an underexplored disanalogy between the moral and epistemological domains. While sentimentalism has a rich tradition as a view about the nature of moral judgment,¹⁶ analogous views about epistemological judgment seem much less promising, and have never (as far as I know) been defended in print. This disanalogy is important because it means that certain argumentative strategies that are available in metaethics—strategies that involve appealing to the sentiments—will be unavailable in metaepistemology.¹⁷

Even if we can't explain epistemic akrasia in terms of epistemic emotions, however, we might still be able to provide an account of epistemic akrasia in terms of inner conflict. Allan Gibbard offers an account of akrasia that doesn't depend on (though it is compatible with) the notion of an emotional motivational system:

This is a picture of two motivational systems in conflict. One system is of a kind we think peculiar to human beings; it works through a person's accepting norms. We might call this kind of motivation *normative* mo-

¹⁵See, e.g., Haidt (2003).

¹⁶See Hume (2007). Prinz (2007) offers a more recent development of the view.

¹⁷This isn't to say that sentiments are never directed at beliefs. The following possibilities are completely familiar: I might feel guilty for doubting that my friend will keep her promise; I might be angry at you for believing that I won't keep *my* promise; a fundamentalist or ideologue might feel blame or contempt for others who fail to share her beliefs. These examples suggest that, insofar as we are metaethical sentimentalists, we should allow that it is possible to make moral judgments not only about actions, but also about beliefs. That is, because the range of negative(positive) sentiments that are associated with moral disapproval(approval) can be directed at beliefs as well as actions, metaethical sentimentalists should allow that just as one can judge that it is morally wrong(right) to act in certain ways, one can also judge that it is morally wrong(right) to hold certain beliefs.

None of this, however, should encourage somebody hoping to give a sentimental account of distinctively *epistemological* judgments. In fact, it just makes the difficulty of that project more apparent—it's hard to see how a sentimentalist account of normative judgments about beliefs won't end up just telling us what it is to judge that beliefs are *morally* (as opposed to epistemically) permitted, required, or forbidden.

tivation, and the putative psychological faculty involved the *normative control system*. The other putative system we might call the *animal control system*, since it, we think, is part of the motivational system that we share with the beasts. Let us treat this picture as a vague psychological hypothesis about what is going on in typical cases of “weakness of will.” (1990, p.56)

In the absence of more of a story about what distinguishes these motivational systems, we might find this answer unsatisfying. However, Gibbard has more to say on the topic:

In many cases a norm is stated in language, or thought in words. A norm, we might say, is a linguistically encoded precept. Perhaps, then, we should think of the motivation I have been calling “normative” as motivation of a particular, linguistically infused kind—a kind of motivation that evolved because of the advantages of coordination and planning through language. (1990, p.57)

While Gibbard is offering an account of moral/practical akrasia, his account generalizes naturally to epistemic akrasia. The notion of a linguistically infused belief system that operates alongside a non-linguistic belief system seems just as applicable in the epistemic case as the notion of competing motivational systems is in the moral/practical case. For example, imagine a gambler who is disposed to bet on red if the last 10 spins of the roulette wheel have come up black, unless she says to herself something like the following: “The gambler’s fallacy is a fallacy, and the fact that black has come up many times recently doesn’t mean that red is any more likely to come up the next time.” Suppose that when she does remind herself of these facts, she bets in line with the true probabilities, rather than the ones that would be suggested by the gambler’s fallacy.

In such a case, we might distinguish between the dispositions produced by her non-linguistic belief system—call these “beliefs_n”—and the dispositions produced by her linguistically infused belief system—call these “beliefs_l”. We might go on to say that she believes_n that red is more likely to come up after a string of blacks, but at the same time she believes_l that this is not the case, and that the chances of red coming up on any given spin are independent of the results of the previous spins.

It isn’t obvious that we should describe a case like the above one as a case of epistemic akrasia, since it isn’t obvious that we should describe the agent in the above case as straightforwardly believing that red is more likely to come up after a run of blacks—after all, in some situations (the ones when she reflects) she doesn’t act as if she believes it. However, once we have made a distinction between beliefs_n and beliefs_l, we should allow that in some cases, an agent might be disposed by her linguistically infused system to act as if P , but this disposition might be masked by the operation of her non-linguistic belief system, which we may suppose even more powerfully disposes her to act as if $\sim P$.¹⁸ In such a case, even when she says to herself

¹⁸I intend “masked” to be understood in the sense in which Johnston (1992) raises the possibility of dispositions being masked.

“remember, the gambler’s fallacy is a fallacy” she still acts as if the results of later spins probabilistically depend on the results of earlier ones. Such a case would be a straightforward case of epistemic akrasia.¹⁹ More generally, when one’s linguistic and non-linguistic subsystems are producing conflicting dispositions, we can say that one’s (unsubscripted) beliefs align with the dispositions produced by the system that’s ultimately guiding one’s behavior—if one’s non-linguistic system wins out, then one’s beliefs are her beliefs_n, and if one’s linguistic system wins out, then one’s beliefs are one’s beliefs_l.²⁰

Not much rests on the two systems being distinct *linguistic* and *non-linguistic* belief systems. As long as we can identify *some* distinct psychological subsystems, each of which has a right to be called a belief system, then we can understand cases of epistemic akrasia as cases in which the outputs of these distinct subsystems conflict. For instance, we might appeal to dual process theories in psychology and the “system 1/system 2” distinction they invoke (Stanovich and West, 2000). Alternatively, Tamar Gendler (2008) has drawn a distinction between belief and “alief” that could play a similar role in grounding an inner-conflict account of epistemic akrasia. Daniel Dennett (1978) draws a distinction between belief and opinion in “How to Change Your Mind,” that might also do the trick. I’ll continue to use the labels “linguistic” and “non-linguistic,” and the subscripts “*l*” and “*n*” to describe the potentially conflicting belief systems, but nothing in what follows depends on these labels being appropriate.

I warned earlier that the account of epistemic akrasia I’d sketch would not be metaepistemologically neutral. In particular, the account I’ve sketched sits nicely with a more general story about epistemological beliefs that might be called “expressivist.” Metaethical expressivists try to understand moral beliefs by thinking of them not as regular beliefs with a distinctive sort of content (namely, content concerning morality), but instead in some other way (often, e.g., by treating them as constituted by non-cognitive motivational attitudes).²¹ The account of epistemic akrasia I’ve sketched suggests a structurally similar strategy in the epistemological case.²²

¹⁹Well, a *relatively* straightforward case, subject to the qualification mentioned in footnote 3.

²⁰This account of epistemic akrasia isn’t completely neutral with respect to accounts of ordinary, non-epistemological belief. In particular, views of belief on which all beliefs essentially involve language, and in which animals properly speaking don’t have beliefs, have a more difficult time allowing for a distinction between a linguistically infused belief system and a non-linguistic belief system. See, e.g., Davidson (1982). Views on which there are constitutive connections between belief and action generally, rather than belief and specifically linguistic action, have an easier time drawing such a distinction. See, e.g., Stalnaker (1984) and Dennett (1987a). I set this worry aside in what follows.

²¹See Blackburn (1993, 1998), Gibbard (1990, 2003). I don’t mean to imply that Blackburn and Gibbard take expressivism to be plausible only as an account of moral judgment; both of them are quite explicit in hoping to give broadly expressivist accounts of epistemological judgment as well, even if they devote more attention to the moral case. I also don’t mean to hold that expressivists must deny that we in fact have moral beliefs—many expressivists (including Blackburn, and more recently Gibbard) are happy to say (roughly) that any sentence in the indicative mood can express a belief, though they go on to hold that belief is a less psychologically unified category than we might have thought, and that normative beliefs differ in significant ways from non-normative beliefs.

²²I say merely “structurally” similar because the strategy does not involve treating epistemological

That is, rather than understanding epistemological beliefs as having a distinctive sort of content, we could try to understand them in some other way.²³ In particular, we might understand them as distinctive in terms of the psychological mechanisms that produce and maintain them, rather than as distinctive in terms of their subject matter. The belief that my evidence supports the claim that I have hands, then, could be understood as having the same content or subject matter as the belief that I have hands; it would be distinctive only in being produced or maintained by a distinctive sort of psychological system (e.g., a linguistically infused system).

Metaethical expressivists often appeal to some version of “judgment internalism”—the view that there is some kind of necessary connection between judging that one ought to act in particular way, and being motivated to act in that way—to support their expressivism.²⁴ The sort of expressivist strategy I’ve just sketched dovetails nicely with an analogous internalist view concerning epistemological judgment, to the effect that there is a necessary connection between judging that one ought to hold a particular belief, and being disposed to hold that belief. On the view I’ve sketched, the judgment that one ought to believe that *P* is in effect a *species* of belief that *P*—it is a sort of belief that *P* that is maintained by a particular sort of psychological system.

Whatever one thinks about judgment internalism in *metaethics*, I take the analogous view to be extremely plausible in *metaepistemology*, and I hope the following digression will make it clear why. Metaethical judgment internalists face challenge of accounting for the apparent possibility of the amoralist—someone who makes moral judgments, but is not at all disposed to act in accordance with them.²⁵ Fiction abounds with apparent amoralists; “I know what *morality* says I should do, but I don’t care about morality!” has a good deal of resonance. While judgment internalists have long been aware of this challenge and have strategies for answering it, it’s controversial just how successful those strategies are.²⁶ But metaepistemological judgment internalists face no comparable challenge in accounting for the apparent possibility of the “a-epistemicist”—someone who makes epistemological judgments but is not at all disposed to believe in accordance with them—for there is no such apparent possibility.²⁷

What might be the rallying cry of such an a-epistemicist? Perhaps the following: “I

beliefs as non-cognitive, or desire-like.

²³In recent years, many writers have explored expressivist strategies for understanding epistemological judgments, especially when “expressivist” is read in the broad way that I suggest above. See, e.g., the papers presented at the recent “Epistemic Expressivism Workshop” at the University of Edinburgh.

²⁴The term “judgment internalism” was coined by Darwall (1983), though the strategy of using such theses to support broadly expressivist views in metaethics goes back a good deal further.

²⁵See Brink (1989) for a paradigmatic statement of this challenge.

²⁶See Blackburn (1998) for a nuanced treatment of the challenge of amoralism as manifested in the character of Milton’s Satan. Blackburn’s treatment is ultimately sympathetic to broadly judgment internalist views.

²⁷I apologize for the terrible name. But try as I might, I can’t think of a better one for the character who is to metaepistemology what the amoralist is to metaethics. Luckily, if I’m right that there is no serious challenge posed by such a character, we don’t need to worry too much about what to call her.

know what the *evidence* suggests is true, but evidence is a systematically unreliable guide to the truth!” But this won’t do; plausibly, to think that some source of information is a systematically poor guide to the truth of whether *P* is just to think that it doesn’t provide good *evidence* of whether *P*. The rallying cry of the a-epistemicist, then, could only be uttered by somebody who was deeply confused about what “evidence” means.²⁸ While broadly expressivist views in metaethics are often taken to be more plausible than their analogs in metaepistemology, I take considerations like those above to put at least some pressure on this idea.²⁹ One influential line of argument for expressivism in metaethics—that which appeals to judgment internalism—is if anything *more* powerful when used to motivate analogous views in metaepistemology.

Much more would have to be said to develop a distinctive account of epistemological judgment based on what I’ve said so far, and it would be beyond the scope of this dissertation to do so. However, I do want to show how one *prima facie* problem for such a strategy—in particular, for the claim that epistemological beliefs do not have a distinctive subject matter—can be addressed compatibly with everything I’ve said so far.

The apparent problem I want to deal with concerns the possibility of justified false belief. An example will bring out the worry. Suppose you believe (non-akratically) that your car is parked in your garage, where you left it. As it turns out, however, I have recently stolen your car and sold it to my neighborhood chop-shop. Now, consider what *I* should think about the following two beliefs of yours:

- Your car is parked in your garage.
- You ought to believe that your car is parked in your garage.

It seems that I should think that the first belief of yours false, while the second belief is true—I should take you to have a justified false belief.³⁰ But (and this is the puzzle) if I am to coherently think this, then it seems as if your two beliefs must concern different subject matters—they can’t both just be beliefs about where your car is parked, because then there would be no way for one of them to be correct while the other is incorrect. Is there any way we can account for this case without positing a distinctive epistemological subject matter for you to be getting things right about (even though you’re getting things wrong with respect to the subject matter of where your car is parked)? More generally, how can we account for the possibility of justified false belief without positing a distinctive epistemological subject matter?

In general, expressivists adopt the strategy of explaining normativity indirectly, by explaining what is involved in making normative judgments. This case is no

²⁸It won’t do to point to the possibility of skepticism about the reliability of reason; someone who thinks, e.g., that careful reflection is a poor guide to the truth is not someone who thinks that *evidence* is a poor guide to truth, but instead someone who thinks that the fact that a view seems plausible after careful reflection is at best weak evidence that it is true.

²⁹Cuneo (2007), for one, takes realism to be more plausible in metaepistemology than metaethics, and pursues the dialectical strategy of defending realism in metaethics by arguing that it stands or falls with realism in metaepistemology.

³⁰Assuming, at least, that you had no reason to expect me to steal your car.

different—the epistemic expressivist should understand her burden as the burden not of directly explaining the possibility of justified false belief, but instead as the burden of explaining what it is involved in thinking that somebody has a justified false belief.³¹

Cases in which it is attractive to think of some situation that it is one in which someone has a justified false belief are typically cases in which we are thinking of the situation in a third-personal sense—cases in which the subject of the situation is not me-here-now, but some other person, or me at some other time. In the example above, it is *I* who thinks that *you* have a justified false belief, and this example is typical in the relevant respects. What we need to account for, then, is the possibility of believing claims of the following sort:

- She (not me) ought to believe that P , but in fact $\sim P$.

Nothing I've said so far straightforwardly bears on the question of what is involved in having such beliefs. However, it's plausible that just as an expressivist account of what it is for me to believe that I-here-now ought to believe that P will turn on my current beliefs_{*i*} about whether P , an expressivist account of what it is for me to believe that someone else ought to believe that P will turn on something like my conditional beliefs_{*i*} about P , given only information available to the relevant person. That is, I might think that P , while thinking that some body of evidence E (not *my* evidence) supports believing that $\sim P$. This might involve believing_{*i*} that P , while having a kind of conditional belief_{*i*} in $\sim P$ given E .

Of course, this suggestion is only helpful insofar as we have some grip on the idea of a conditional belief_{*i*}, where the condition is some body of information other than my own. But I don't think this is such an unusual or strange idea. Conditional beliefs are relatively familiar (even if their nature is not entirely transparent), and the same considerations that lead us to distinguish between belief_{*i*} and belief_{*n*} should lead us to recognize similar distinctions among conditional beliefs. For instance, in the case of MATT, Matt believes_{*i*} that flying isn't particularly dangerous, while he believes_{*n*} that it is. We should say that similar distinctions arise in Matt's conditional beliefs—for instance, Matt may fail to hold the conditional belief_{*i*} that he'll die tomorrow, given that he'll fly tomorrow, even while he does hold the conditional belief_{*n*} that he'll die given that he'll fly.

The distinction between beliefs_{*i*} and beliefs_{*n*} is less central, however, to the explanation of what's involved in thinking that someone *else* has a justified false belief than it is to the explanation of what's involved in thinking of *oneself* that one has a justified false belief (i.e., being epistemically akratic). On the view I've been exploring, epistemic akrasia *requires* a conflict between one's beliefs_{*i*} and one's beliefs_{*n*}, but thinking that somebody else has a justified false belief does not—even if my conditionals beliefs_{*i*} and my conditional beliefs_{*n*} are in sync, I might believe that someone else has a justified false belief, so long as my *actual* beliefs diverge in the appropriate way from my conditional beliefs.

³¹Which is not to say that she might not also agree that justified false beliefs are possible. It's just that in doing so, she'll be making substantive epistemological claims, rather than sticking to metaepistemology.

Even if the above strategy suffices to ward off worries about the possibility of justified false belief, there are still many puzzles facing the theorist who tries to explain epistemological judgment without positing a distinctive epistemological subject matter.³² But while I haven't provided anything like a full account of epistemological judgment, I have said enough to shed some light on the question of the rationality of epistemic akrasia, as I'll show in the next section.

2.3 Normative Upshots

I've sketched an account of epistemic akrasia as a species of fragmentation, or inner conflict. Why think, however, that a story about what akrasia involves could be relevant to the question of whether being akratic could be rational? I'll try to answer that question in this section.

When motivating the idea that akrasia could sometimes be rational, I appealed to the claim that the belief that P and the belief that the evidence supports P have different subject matters; the latter claim is about epistemology, while the former is about whatever P is about—aviation, anatomy, or whatever. But I hope it's now clear that this claim isn't so innocent or so obvious as it at first appears. If we accept the account of akrasia sketched in the previous section, then we'll reject the idea that my belief that I ought to believe that P has a different subject matter from my belief that P . Rather, the two beliefs are distinct not in their subject matter, (both of them concern whether P) but in their etiology—they are the products of distinct, competing belief-producing psychological systems. Once we understand things in these terms, understanding akrasia as irrational—in particular, as a species of inconsistency—becomes much more attractive.

Our account holds that being epistemically akratic requires having two competing belief-producing psychological subsystems. Plausibly, if epistemic akrasia were ever rationally required, then ideally rational agents would sometimes be epistemically akratic. This entails that there are some circumstances in which, for an agent to be ideally rational, she would need to have two conflicting belief-producing psychological mechanisms. But it seems implausible that ideal rationality requires anything like that. Indeed, various writers have appealed to the idea that humans are cognitively fragmented to explain how certain forms of *irrationality* are possible for us. (Lewis, 1982; Stalnaker, 1984, 1991, 1999). An agent with a single belief-producing psychological system could, plausibly, be ideally rational, so long as the beliefs produced by that system obeyed certain constraints (e.g., they would have to cohere with one

³²For instance, it would be hasty to simply *identify* one's beliefs about what one ought to believe with one's beliefs_{*i*}. We might, for instance, allow that one can believe_{*i*} that P without believing that one ought to believe that P , if one thinks that one's case is a *permissive* one—that is, if one thinks that one's evidence permits one either to believe P , or to believe $\sim P$. White (2005) argues against the possibility of such cases, but even if we agree with him, we might want to have some sense of what it is to *think* that some case is a permissive one. I take the question of what's involved in taking a permissive attitude towards one's beliefs to be an interesting, tricky one. I suspect that answering it will require saying something about how one interacts with others.

another in certain ways, be appropriately responsive to information from the agent's perceptual system, and so on).

If the belief that one ought to believe that P were about a distinctive normative subject matter—one independent of the subject matter of P —then there would be no *prima facie* obstacle to the possibility of rational epistemic akrasia. One might have reasons that indicated that the correctness condition of the belief that P was satisfied, while at the same time having reasons that indicated that the correctness condition of the belief that one ought to believe that P was not satisfied. But if we accept the account of akrasia as fragmentation, this option isn't available. If akrasia is fragmentation, then being akratic amounts to having two inconsistent beliefs (a belief _{i} and a belief _{n}) about the same subject matter, rather than having two consistent beliefs concerning distinct subject matters.

Of course, none of this settles the issue of the rationality of epistemic akrasia. After all, some writers have argued that inconsistent beliefs and cognitive fragmentation might sometimes merit positive epistemic appraisal.³³ Certainly if we're moved by the idea that ought implies can in epistemology, we might allow that epistemic akrasia is sometimes rational on the grounds that cognitive fragmentation is an inescapable aspect of the human condition.³⁴

Still, the fragmentation account of akrasia changes the debate about the rationality of akrasia, even if it doesn't settle it for one side or the other. Before, the idea that akrasia might sometimes be rational looked to be a straightforward consequence of an innocent claim: the claim our evidence concerning a particular subject matter—the matter of what we ought to believe—can be misleading. This claim seemed to follow from a more general claim, also innocent, to the effect that justified false belief is possible in any domain; for any subject matter, our evidence can sometimes be misleading with respect to that subject matter. Now, these *prima facie* innocent claims are revealed as carrying with them substantive ideological presuppositions, at least when applied to the case of epistemic akrasia—in particular, they presuppose that epistemological beliefs should be understood as having a distinctive (epistemological) subject matter, which we've seen is an eminently contestable notion. Moreover, to the extent that the view that akrasia is irrational turns out to be closely related to the idea that logical consistency is an epistemic ideal—on the fragmentation account

³³For the former, see Christensen (2007). For the latter, see Egan (2008). Though importantly, neither argues that being rationally ideal in all respects is compatible with inconsistent or fragmented belief. Christensen argues that rational ideals sometimes conflict, and that respecting certain ideals will sometimes require having inconsistent beliefs. Egan argues that for agents with unreliable belief-forming modules, being cognitively fragmented may be an effective form of "damage control,"—one that will stop the beliefs formed by the unreliable modules from spreading to the rest of the agent's cognitive system.

³⁴A number of writers have discussed the principle that ought implies can specifically in the context of epistemology—some endorsing it, some attacking it. Fred Dretske (2000, p.598) endorses a strong epistemological ought implies can principle, and uses it to argue that we have a right to hold perceptual beliefs because we are psychologically incapable of failing to form them. William Lycan (1985, p.146) suggests that ought implies can is less plausible in epistemology than in ethics, and expresses doubt that beliefs can be warranted in virtue of our being psychologically incapable of abandoning them. My sympathies are closer to Lycan's than to Dretske's. See Greco (Forthcoming).

of epistemic akrasia, epistemic akrasia is really a species of logical inconsistency—we may find it more plausible as a result.³⁵

A different normative (or metanormative?) conclusion one might draw from the fragmentation account of akrasia is that asking about the rationality of epistemic akrasia is a species of category mistake. For instance, we might think of epistemic akrasia as involving a conflict between two belief-like states (or species of belief), but think that only one of those states is rationally evaluable. We could support this position by appealing to Gendler's distinction between belief and "alief." In Gendler's framework, aliefs are "so-called because alief is associative, automatic, and arational." (2008, p.641). The key word for my purposes is "arational." If we think of epistemic akrasia as involving something like a conflict between belief and alief, then to ask whether akrasia might sometimes be the rationally appropriate response to unusual evidence would be to make a category mistake. It would be like asking whether it might sometimes be rationally appropriate to both believe that $2 + 2 = 4$ and be tall. While we can ask about the rationality of mathematical beliefs, asking about the rationality of being tall is confused. If being epistemically akratic is relevantly like both believing that $2 + 2 = 4$ and being tall—if it amounts to being in two states, only one of which is rationally evaluable—then it's a mistake to ask whether epistemic akrasia might sometimes be rationally appropriate. Rather, we should just ask about the rationality of belief, and should set aside questions about the rationality of akrasia (which involves alief) as resting on false presuppositions. While this wouldn't quite amount to a vindication of the idea that epistemic akrasia is irrational, it *would* amount to a rejection of the positive arguments, considered in §1.2, that epistemic akrasia is sometimes rationally appropriate.

My suspicion is that some cases that it's tempting to describe as cases of epistemic akrasia can be happily thought of as involving conflicts between two states, only one of which is rationally evaluable. In such cases, the verdict described above will be appropriate; asking about the rationality of being in such combined states will amount to making a category mistake. I suspect that other cases of epistemic akrasia aren't so happily thought of in these terms; in particular, I suspect that other cases of epistemic akrasia are more comfortably thought of as closely related to straightforwardly inconsistent belief, which we are usually prepared to describe as rationally evaluable (in particular, we are usually prepared to describe straightforwardly inconsistent belief as irrational). In such cases, we needn't hold that questions about the rationality of epistemic akrasia are confused.

While drawing such distinctions might seem awkward, I do not assume that the notion of epistemic akrasia cuts at the psychological joints, or that every case that it's tempting to describe as a case of epistemic akrasia must ultimately be given the same treatment. There is a spectrum of cases, all of which might be described as involving something like epistemic akrasia, and which vary in the degree to which rational evaluation seems appropriate. Start with somebody who gets extremely anxious when flying, but doesn't let this anxiety affect her behavior in any substantial way (she doesn't, e.g., avoid flying when the stakes are high). While we *might* describe such

³⁵Or perhaps less plausible, depending on our views about the putative ideal of logical consistency.

a subject as irrational for experiencing such anxiety—perhaps by interpreting such anxiety as a manifestation of an irrational belief in the dangers of flying—it strikes me as much more natural to treat her anxiety as *arational*. On the other end of the spectrum we have cases like that of MATT, discussed at the beginning of this essay. Since Matt’s aversion to flying *does* influence his behavior in substantial ways, even when the stakes are high, it’s attractive (at least to me) to interpret Matt as having an irrational belief that flying is dangerous. At the very least, Matt seems more open to rational criticism than someone who merely gets anxious when flying, but who still flies when she has strong reasons to do so. But Matt’s case plausibly differs only by degree from that of someone who is merely anxious when flying.

In this section I’ve leaned heavily on the idea that if the fragmentation account of epistemic akrasia is correct, then epistemological beliefs don’t have a distinctive subject matter, and I’ve used this idea to defuse some of the motivation for thinking that epistemic akrasia must sometimes be rational. But I fear that in leaning so heavily on this idea, I may merely undermine the support for the fragmentation account of akrasia, rather than changing any minds about the rationality of akrasia. So in the next section, I’ll try to provide some independent support for the fragmentation account of akrasia (and the “no distinctive epistemological subject matter” claim that goes with it) by showing how it treats putative cases of “higher-order” akrasia.

2.4 Higher-Order Akrasia

So far we’ve just discussed the difference between first-order beliefs and beliefs about what you ought to believe, but at least formally, we can replicate this distinction at arbitrarily high orders. That is, we can also distinguish between beliefs about what you ought to believe and beliefs about what you ought to believe you ought to believe, and so on. But should we? Consider the following series, which we could continue indefinitely:

1. The belief that P
2. The belief that one ought to believe that P
3. The belief that one ought to believe that one ought to believe that P
4. The belief that one ought to believe that one ought to believe that one ought to believe that P

Once we grant that items 1 and 2 are distinct in their content (rather than merely being distinct in their etiology, as the fragmentation account of akrasia suggests), it’s natural to say the same thing about items 2 and 3, and items 3 and 4, and so on. We’ll hold that with each step up, we get a belief about a new epistemological proposition that one might fail to hold, even while holding the belief one level earlier than it on the list. We might call cases in which one’s level 1 and level 2 beliefs come apart “first-order” akrasia, cases in which one’s level 2 and level 3 beliefs come apart “second-order” akrasia, and so on. And as a matter of fact, the best worked out

views on which first-order akrasia is sometimes rational also entail that higher-order akratic states are sometimes rational.³⁶

But before the question of whether higher-order akrasia might sometimes be rational even arises, we might ask whether it's possible at all. After all, while it's easy to come up with examples (like MATT) that are naturally described as examples of first-order akrasia, it's hard to continue the process even just one level up, let alone two or three levels up. I'm not at all sure what it would be for someone to be second-order akratic, or how somebody who had such beliefs would be disposed to act. Of course we can imagine somebody being disposed to utter the sentence "I ought to believe that *P*, but I oughtn't believe that I ought to believe that *P*," and closely related sentences, but that's about it. This is suspicious; normally beliefs are capable of being manifested in more than one way.³⁷

One might respond that while states of higher-order akrasia are possible, such states are psychologically complex, and it's hard for us to conceive them. There's some precedent for this sort of move; there are grammatical facts about sentences with complex embeddings, but we often find it hard to recognize such facts due to the difficulty of parsing such sentences. But the response is ultimately unpersuasive. We often *can* make sense of higher-order psychological states that seem just as complex as states of second and third-order akrasia. To take an example of Jane Heal's, "I may be furious that you are amused at my unease at suspecting that you have realized that I have been gloating over my supposed exclusive knowledge of some unsavoury habit of yours." (1978, p. 120) While the sentence takes a bit of effort to parse, once we've done so, we can have a reasonably clear picture of what the state of mind Heal describes would involve, what might cause someone to be in such a state, and how it could be manifested in a variety of behaviors (not just the disposition to utter a few specific sentences).³⁸

³⁶I have in mind Williamson's (2000) views on higher-order evidential probability. According to Williamson, just as *P* can be probable while the proposition *that P is probable* can be improbable, the proposition *that P is probable* can be probable while the proposition *that P is probable is probable* can be improbable.

³⁷See Schwitzgebel (2002), who defends an account of belief on which believing that *P* is a matter of having a sufficiently similar set of dispositions to those of a "stereotypical" *P*-believer, and in which there will be a family of behavioral, phenomenal, and cognitive dispositions stereotypically associated with any given belief. I'm suggesting that in the case of higher-order akrasia, it's very hard to identify any such stereotypical dispositions beyond a very narrow array of linguistic dispositions.

³⁸Another lesson we can learn from Heal's example is that our inability to make sense of certain higher-order *intrapersonal* distinctions (e.g., the distinction between my believing that I ought to believe that *P*, and my believing that I ought to believe that I ought to believe that *P*) should *not* lead us to doubt the importance of analogous higher-order *interpersonal* distinctions. There are clearly important differences between my believing that my evidence supports the claim that *P*, and my believing that my evidence supports the claim that *your* evidence supports the claim that *P*. In particular, accounting for distinctions among higher-order *interpersonal* doxastic/epistemic state of this sort doesn't seem to require positing more than a single belief system; in the *interpersonal* case, beliefs at the different levels have straightforwardly different behavioral consequences, even when we confine our attention to a single belief producing system (whether linguistically infused or not). Relatedly, there's no temptation to describe the state of believing that *P* while also believing that *someone else's* evidence doesn't support the claim that *P* as involving akrasia or inner conflict of any sort.

The fact that there don't seem to be distinctive behavioral manifestations of higher-order akratic states (other than a narrow range of sentences they might dispose one to utter) does not show that such states don't exist, or that there is no distinction to be made between first-order and higher-order akrasia.³⁹ But it is suggestive, and in the presence of an independently well-motivated account of epistemic akrasia that doesn't make room for a distinction between first-order and higher-order akrasia, we might reasonably take our inability to make intuitive sense of such a distinction as further evidence for its non-existence.

As a matter of fact, the fragmentation account of akrasia turns out to mesh nicely with the denial of the distinction between first-order and higher-order akrasia. If we understand first-order epistemic akrasia as involving a state of conflict between two competing belief-producing psychological subsystems, then it's natural to understand second-order akrasia as involving yet another *distinct* belief-producing psychological subsystem which might compete with the original two belief-producing psychological subsystems. Second-order akrasia would involve one's linguistic belief system disposing one to believe that *P*, and this third, distinct belief system disposing one not to believe that *P*. Third-order akrasia would require a fourth belief-producing psychological mechanism, and so on all the way up; accepting the possibility of epistemic akrasia of arbitrarily high orders would require accepting the existence of infinitely many distinct psychological subsystems, each of which produces belief-like states. Since we obviously do not have an infinite hierarchy of distinct belief-producing psychological subsystems, the fragmentation account of epistemic akrasia is not compatible our being potentially akratic at arbitrarily high orders.

The difficulties associated with making sense of akrasia of arbitrarily high orders go beyond the contingent fact that we lack an infinite array of distinct belief-producing psychological subsystems. Even if we had many such subsystems, it isn't clear what features they would need to have to be appropriately placed in the sort of hierarchy they'd need to be placed in for there to be a distinction between second-order akrasia, third-order akrasia, and so on. Earlier I suggested that the belief system associated with non-epistemological beliefs might not rely on language, or might work quickly and automatically, while the belief system associated with epistemological beliefs might crucially involve language, or perhaps might be triggered by reflection and conscious deliberation. If we are sympathetic to this idea, then we should be able to make sense of first-order akrasia—it will involve a belief system with certain characteristic features (non-linguistic, fast, automatic) disposing one to believe one way, and another belief system with different characteristic features (linguistic, slow, consciously accessible) disposing one to believe another way, where the former system wins out.

³⁹Should we hold that there's no difference between, e.g., being a three-dimensionalist and being a four-dimensionalist on the grounds that the difference only seems capable of manifesting itself verbally, and then only in a narrow range of situations? Maybe, but this would be a highly controversial position that would require rejecting a wide range of metaphysical debates as merely verbal. While some contemporary anti-metaphysicians do seem to take something like the above position, nothing so strong is required for the line I'm defending here.

But even if we can make sense of first-order akrasia, we don't have a way of making sense of second-order akrasia. Presumably, it would involve a belief system with certain characteristic features (linguistic, slow, conscious) disposing one to believe one way, and another belief system with different characteristic features (???) disposing one to believe another way, where the former system wins out. But it's completely obscure what the features characteristic of this latter system would need to be. For a creature to be potentially akratic at arbitrarily high orders, it wouldn't just need infinitely many distinct belief systems—there would also need to be a well-motivated way of hierarchically ordering this infinitude of belief systems by their distinctive features such that each order was associated with a belief system with certain distinctive features.

Since I doubt that there is such a well-motivated way of hierarchically ordering distinct possible belief-producing subsystems, I doubt that the fragmentation account of akrasia can generalize to give us an account of epistemic akrasia of higher-orders. I take this to be a welcome result for that account, since we don't seem to be able to make sense of distinctions between different states of higher-order akrasia.

Let's take a step back. In this section, I've tried to provide some independent support for the fragmentation account of epistemic akrasia, and the associated claim that the belief that P and the belief that one ought to believe that P differ in their etiology, rather than in their subject matter. I argued that if the two beliefs differ in their subject matter, then we should be able to make sense not just of first-order akrasia, but of second-order akrasia, third-order akrasia, and so on. But we cannot make sense of such states. In the absence of a principled explanation of why first-order akrasia should be possible even if higher-order akrasia is not, we might conclude that even though there *are* distinctions between different higher-order akratic states, our ability to grasp those distinctions is extremely limited. But the fragmentation account of akrasia provides a principled reason for denying that there are distinctions to be made between distinct higher-order akratic states; such distinctions would require a hierarchically ordered infinitude of belief-producing psychological systems, but such an infinitude does not exist.

2.5 Conclusion

I started this chapter by presenting a puzzle. On the one hand, there are strong temptations to deny that epistemic akrasia could ever be rational—it seems like a species of incoherence or inconsistency.⁴⁰ On the other hand, if we can have misleading evidence about a particular subject matter—what we ought to believe—then it seems as if appropriately responding to that evidence will require being epistemically akratic. I argued that we can resolve this puzzle by getting clearer about what sort of psychological state epistemic akrasia is. Once we understand it as a species of fragmentation, we can hold that it is always irrational, while explaining away the

⁴⁰Though in light of the observations at the end of the considerations at the end of §3, we might in some cases stop short of calling epistemic akrasia *irrational*, and instead hold that it is sometimes merely *arational*.

appeal of arguments that suggest otherwise. These arguments rely on a false presupposition about the nature of epistemological beliefs—the presupposition that they concern some distinctive subject matter about which we might have misleading evidence. Once we reject this presupposition in favor of the fragmentation account of akrasia, the claim that akrasia is always irrational looks much more innocent; a rational agent is never at war with herself.

Chapter 3

Verbal Debates in Epistemology

3.1 Introduction

If you're a professional philosopher, and you're honest with yourself, you've probably worried at some point or another that some debate in which you have a stake is ultimately merely verbal, or terminological, or somehow non-substantive. You've almost certainly thought that some debates in which *other* philosophers have stakes are merely verbal. This suspicion has historically arisen most often in the context of metaphysics. Hume (1740/2006) told us that metaphysics can contain nothing but sophistry and illusion, and logical positivists like Carnap (1950) were of a similar mind. While the particular commitments that led Hume and Carnap to these conclusions are now much less popular, skepticism about whether (at least some) metaphysical questions are substantive is alive and well. There's now an entire sub-field of philosophy—metametaphysics—that consists largely of debates about whether certain metaphysical questions might somehow be non-substantive or merely verbal.¹

I won't try to resolve any disputes about whether particular metaphysical questions are or are not substantive in this chapter. Rather, I'll explore analogous issues concerning epistemological questions. I'll organize my discussion around two challenges to the substantiveness of certain epistemological questions. While they seem quite different *prima facie*, they are in fact closely related to one another. The first challenge derives from the work of William Alston, and the second from the burgeoning field of experimental philosophy. I'll argue that, while these critiques may have some limited success, they cannot serve as wedge cases for a more wide-ranging skepticism about the substantiveness of epistemological debates; most epistemological questions are immune to the worries they raise.

Before getting into the details, however, it will help to have an example of a question that is plausibly merely verbal, and to say a bit about what I'm assuming about the distinction between substantive questions and merely verbal ones. In his lecture "What Pragmatism Means," William James discusses a debate he once observed that

¹See the essays collected in Chalmers et al. (2009). In this chapter I'll use 'non-substantive' and 'merely verbal' as synonyms. I'll call questions that are not merely verbal 'substantive.' As I mention later, I do not assume that this distinction is a sharp one.

all of us, I hope, can agree is in some sense merely verbal:²

Some years ago, being with a camping party in the mountains, I returned from a solitary ramble to find every one engaged in a ferocious metaphysical dispute. The corpus of the dispute was a squirrel – a live squirrel supposed to be clinging to one side of a tree-trunk; while over against the tree's opposite side a human being was imagined to stand. This human witness tries to get sight of the squirrel by moving rapidly round the tree, but no matter how fast he goes, the squirrel moves as fast in the opposite direction, and always keeps the tree between himself and the man, so that never a glimpse of him is caught. The resultant metaphysical problem now is this: *Does the man go round the squirrel or not?* He goes round the tree, sure enough, and the squirrel is on the tree; but does he go round the squirrel? In the unlimited leisure of the wilderness, discussion had been worn threadbare. (James, 1948, p. 141)

James goes on to give a particular account of *why* the question of whether the man goes round the squirrel or not is a merely verbal question; James' view is (roughly) that questions are substantive only when they have some practical import—only when deciding them one way rather would make a difference to how one would behave: “What difference would it practically make to any one if this notion rather than that notion were true? If no practical difference whatever can be traced, then the alternatives mean practically the same thing, and all dispute is idle. Whenever a dispute is serious, we ought to be able to show some practical difference that must follow from one side or the other's being right.” (1948, p.142) Because we could regiment our use of the word ‘round’ in various ways, none of which would lead to our behaving differently, James regards the question of whether the man goes round the squirrel as merely verbal.³

However, we needn't accept this account of what it is for a question to be merely verbal in order to think that there's *some* distinction in the neighborhood, and that the question in James' anecdote falls on one side of that distinction. For instance, the question of whether there is an odd or even number of stars in the universe (assuming there are finitely many), while probably of no practical importance, is plausibly a substantive matter—at the very least, unlike the question James discusses, it's not the sort of question that seems to call for linguistic regimentation.⁴

I don't have an alternative account of the distinction merely verbal questions and substantive ones, and I don't assume that some reductive account can be had.⁵

²David Chalmers (Forthcoming) also uses James' case as a paradigm example of a merely verbal dispute.

³Presumably James means to exclude merely linguistic behavior.

⁴Or maybe it does—perhaps ‘star’ is vague, and different precisifications of ‘star’ would lead to different answers to the question of whether there are an odd or even number of stars in the universe. But in that case, we can focus on the question of whether there are an odd or even number of stars', where ‘star’ is some acceptable precisification of ‘star.’ Our question will still be of no practical importance, but will not be merely verbal in any interesting sense.

⁵For some attempts to give informative accounts about what the distinction amounts to, see the essays by Hirsch and Chalmers in Chalmers et al. (2009).

Moreover, for familiar reasons, we should be doubtful that the distinction between substantive questions and merely verbal ones is a precise one. It will admit of borderline cases, and it won't always be clear whether some question calls for a determination of fact or a regimentation of linguistic usage.⁶ Vague distinctions are still distinctions, however; even after we admit that the distinction between substantive and merely verbal questions is a vague one, and even if we despair of finding a general reductive account of the distinction, we might still reasonably ask about the status of particular questions. The question in James' anecdote, I'd suggest, is a paradigm case of a merely verbal question, and we might wonder whether particular epistemological questions are relevantly (dis)similar to it without assuming that the sense of (dis)similarity at issue can be precisely or reductively characterized. In particular, I'll argue in §3.2 that a wide range of epistemological questions are *clear* cases of substantive questions; it's compatible with my arguments that some other cases may be harder to classify.

While particular views about the distinction between substantive and merely verbal questions can be quite controversial, I hope that it is not controversial that there is *some* legitimate distinction in the area. Once we accept this distinction, why might we worry that certain epistemological questions are merely verbal?

3.2 Two Challenges

3.2.1 Alston on Justification

In "Epistemic Desiderata" (1993), William Alston argued that many questions about what's required for a belief to be justified are not substantive. He later developed this position in his book *Beyond Justification* (2005). Very roughly, Alston's position is that while there are various properties beliefs might have such that it is uncontroversially desirable from an epistemic point of view that beliefs should have these properties, there are no substantive questions about which combinations of these properties beliefs must have to merit the label "justified."

It will help to spell out his position in some more detail to get an idea of which sorts of epistemological questions he thinks are merely verbal. Consider the following list of properties that beliefs can have, each of which has been thought by some philosophers to have some important connection to epistemic justification:⁷

Reliability: A belief might be such that it was produced by an objectively reliable process—a process with a high objective chance of producing true beliefs.⁸

⁶Of course, this issue is closely related to that of the analytic/synthetic distinction. While some philosophers take Quine (1951/1953) to have shown that there is no such distinction, a more moderate (and I think more plausible) position holds that there is a distinction, while conceding that it is a vague one. Johnston (1992), for one, refers repeatedly to the "vagueness of the analytic/synthetic distinction."

⁷This list is based on the lists in Alston's work, specifically (1993, pp. 528-30) and (2005, pp. 12-15). Doubtless many more properties could be added.

⁸See Goldman (1979) for a view identifying justification with reliability in something like the above sense.

Coherence: A belief might be such that it is part of a coherent web of beliefs, in which the various elements of the web bear certain explanatory and inferential connections to one another.⁹

Perceptual Phenomenal Grounding: A belief might be based on perceptual phenomenology. For instance, I might believe that some table is red because it looks red to me.¹⁰

Ratifiability: A belief might be such that its holder would continue to hold it even after sustained reflection on the question of whether it is true.¹¹

Virtuous Basis: A belief might be held as a result of its holder exercising some intellectual virtue.¹²

Defensibility: A belief might be such that its holder is in a position to give a compelling argument in defense of it.¹³

Concerning a similar list of conditions beliefs might satisfy, Alston (1993, p. 531) holds that “for each condition, no one denies that it is desirable to satisfy it, and desirable from an epistemic point of view, desirable vis-à-vis the basic aims of the cognitive enterprise.”¹⁴ Alston goes on to argue that while we can ask, for any particular epistemic desideratum, whether a belief exhibits it, there isn’t a substantive question of which of these desiderata make for justification:

If we take the range of parties to the disputes we have been considering...there does not seem to be enough commonality in their pre-theoretical understanding of the nature of epistemic justification to warrant us in supposing that there is some uniquely identifiable item about which they hold different views. It seems, rather, that they are highlighting, emphasizing, “pushing” different concepts, all called ‘justification’. It seems, to switch

⁹Coherentists such as Quine and Ullian (1978) and early time-slices of Laurence Bonjour (e.g., the Bonjour of *The Structure of Empirical Knowledge*) hold that coherence in something like the above sense is required for justification.

¹⁰See Pollock (1974) and Pryor (2000) for views on which perceptual phenomenology plays an important role in justifying beliefs.

¹¹See Foley (2001) for a view on which something like ratifiability in the above sense is necessary and sufficient for justification.

¹²Virtue epistemologists tend to understand knowledge and justification in terms of epistemic virtues. See DePaul and Zagzebski (2003) for a representative collection.

¹³I’m not sure whether anybody holds that believers must be able to offer actual, explicit arguments in defense of their beliefs for those beliefs to be justified, but this is at least a *possible* view, and one which philosophers often take care to distinguish from their own views. See, e.g., Williamson (2007) and Kelly (2010) on the “dialectical conception of evidence.” Bonjour (1985), however, does seem to think that for a belief to be justified, its holder must have at least some tacit grip on an argument that would vindicate it.

¹⁴Exactly what the idea of a belief’s being “desirable from an epistemic point of view” amounts to is a tricky matter, and it’s not clear how much adverting to the “basic aims of the cognitive enterprise” helps. While Alston thinks the aim of storing true information is basic to the cognitive enterprise, he notes that no parties to debates about justification have been tempted to simply identify justified beliefs with true ones. I’ll avoid these thorny issues here.

to the perspective of this paper, that they are selecting different epistemic desiderata, or packages thereof, as deserving of the honorific ‘justification’.
(1993, p. 534)

The implications of this position are most cleanly brought out when we consider particular cases in which beliefs exhibit some, but not other epistemic desiderata. Consider the following two cases:

Clairvoyant Claire: Claire is a reliable clairvoyant, though she does not believe herself to be clairvoyant; she’s agnostic about the existence of clairvoyance. Claire’s clairvoyance causes her to believe that the president is in New York. While her belief was formed by a reliable process, she doesn’t have any evidence that it is true, nor does she have other beliefs which support the belief that the president is in New York.¹⁵

Hallucinating Hal: It looks to Hal as if there is an apple on the table in front of him, and this leads him to believe that there is an apple on the table in front of him. He takes his vision to be operating normally, and he believes that visual perception is a reliable process. In fact, Hal’s belief is not based on visual perception, but was instead caused by drug-induced hallucination—not a reliable belief-forming process. He has no evidence that this is the case.

Claire’s belief is reliably formed, but it lacks most of the other desiderata mentioned above.¹⁶ Hal’s belief is not reliably formed, but it does seem to have many of the rest of the desiderata on the above list.¹⁷

Epistemologists in the internalist tradition—very roughly, philosophers who hold that whatever justifies our beliefs must be in some sense accessible to our subjective perspective—tend to hold that beliefs like Claire’s are unjustified, while beliefs like Hal’s are justified.¹⁸ Epistemologists in the externalist tradition—very roughly, philosophers who emphasize the epistemological significance of causal/nomological connections between subjects and the external world—tend to give the opposite set of verdicts.¹⁹ According to Alston, however, in asking whether Claire’s and Hal’s beliefs are justified, we’re not really asking a substantive question. Claire’s and Hal’s beliefs each have different things going for them; internalists have correctly identified some desirable features that Claire’s belief lacks which Hal’s belief has, while externalists have equally correctly identified some *other* desirable features that Hal’s belief lacks and which Claire’s belief has. Once we recognize this, according to Alston, we should give up the idea that there’s some further debate to be had between internalists and externalists.

¹⁵This case is inspired by the various cases of reliable clairvoyance discussed by Bonjour (1985).

¹⁶The one tricky case is the “virtuous basis” case, since so much depends on how epistemic virtues are understood.

¹⁷Though for the reasons in the prior note, we might hold that Hal fails to manifest an epistemic virtue in holding his belief.

¹⁸See Bonjour (1985), Conee and Feldman (2004) for two paradigm statements of epistemological internalism.

¹⁹See Armstrong (1973) and Goldman (1979) for two paradigm statements of epistemological externalism.

Fully stating and evaluating Alston's case that debates like the one above—he focuses on debates between internalists and externalists in much of his discussion—are not substantive is beyond the scope of this dissertation. Ultimately, I'll argue that while Alston may be right about the debate over examples like those of Claire and Hal, these are unusual cases; most epistemological debates about which beliefs are justified in which circumstances *can* be vindicated as substantive. While I won't offer a positive argument that debates between internalists and externalists are *not* substantive, the defense I'll offer concerning other debates about justification will not extend to that particular topic. Before getting to my defense, however, I'll introduce another, independent but related challenge to the substantiveness of certain epistemological debates.

3.2.2 The Experimental Critique

In recent years, a number of philosophers have provided empirical support for the claim that there is significant variation in epistemic intuitions along ethnic, socioeconomic, and gender lines.²⁰ To take just one example, Weinberg et al. (2001) presented the following prompt describing a standard Gettier case to their experimental subjects:

Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been stolen, and he is also not aware that Jill has replaced it with a Pontiac, which is a different kind of American car. Does Bob really know that Jill drives an American car, or does he only believe it?

Surprisingly, there were significant cross-cultural differences in the responses.²¹ Western subjects were substantially more likely to answer that Bob only believes that Jill drives an American car, while in East Asian subjects the pattern was reversed.²² Weinberg et al. also found cross-cultural differences in responses to a variant of Keith Lehrer's (2000) "truetemp" case as well as differences along socioeconomic lines in responses to a variant of Fred Dretske's (1970) "cleverly disguised mule" case.

There are a number of ways one might use results like the ones above to provide a critique of standard epistemological practice. One—the one which Weinberg et al. focus on—involves raising a skeptical challenge; insofar as epistemologists (a group composed mainly of high socioeconomic status Westerners) rely on their intuitions in

²⁰Here I'll focus on the work of Weinberg et al. (2001) as a representative sample of work on variation in intuitions across ethnic and socioeconomic lines, though see Buckwalter and Stich (2010) for discussion of variation in epistemic intuitions across gender lines.

²¹The question of whether these results reflect a robust difference is highly controversial. In recent experiments focusing on similar scenarios to the ones studied by Weinberg et al., Nagel (Forthcoming) reports that she "did not find any statistically significant correlations between ethnicity or gender and knowledge ascription."

²²Though the degree to which East Asian subjects were more likely to respond that Bob knows rather than only believes was less than the degree to which Western subjects were more likely to respond that Bob only believes rather than knows.

their practice, we can ask what grounds they have for taking their own intuitions to be more reliable than those of members of other ethnic or socioeconomic groups.²³

However, I want to focus on a different way in which results like Weinberg et al.'s might seem to threaten epistemological practice. One response to such results—sometimes offered as a defense against skeptical challenges²⁴—involves holding that they point to subtle conceptual differences in the different groups surveyed. Rather than speaking of knowledge full stop, in light of such results, perhaps we should speak of various different “knowledge concepts” used by different ethnic and socioeconomic groups. This isn’t an obviously implausible suggestion; when linguists discover variance in grammaticality intuitions along ethnic or socioeconomic lines, a common response is to posit the existence of multiple dialects. Positing a multitude of knowledge concepts in response to studies like that of Weinberg et al. might seem like an equally reasonable response to a similar sort of data.

If the results point to conceptual differences, then we can regard the disagreements between different ethnic and socioeconomic groups as merely apparent. Philosophers (by and large Westerners of comparatively high socioeconomic status) can go on to assert that Bob doesn’t “know” that Jill drives an American car, while granting that East Asians are equally correct when they assert that Bob does “know” this; what they assert is not the same as what we deny. Philosophers can go on constructing theories of knowledge, so long as they understand that such theories will only be adequate to one sense of “knowledge.”

Once we do this, however, we run right back into a version of Alston’s critique. If there are many “knowledge concepts,” questions about whether subjects in various hypothetical circumstances have knowledge or mere true belief will threaten to be merely verbal. How we answer such questions will depend on which knowledge concept we use, and it won’t be clear whether there is anything of substance at stake in the decision to use one knowledge concept rather than another. If we ignore the (putative) fact that there is a multitude of knowledge concepts, the epistemological questions we ask will be of merely parochial interest, as they will be posed using concepts that are the cultural artifacts of rich Westerners. This isn’t *quite* to say that they will be merely verbal; holding fixed which concepts we’re using, there may be substantive questions about whether they apply in particular cases. But this is cold comfort. After all, if we specify fixed a set of rules of application for the notion of “going round the squirrel,” there will be substantive questions about whether the man in James’ anecdote was going round the squirrel, or not. But we may still worry that there’s nothing much of interest in the decision of which set of rules to use, and the associated decision of how to classify cases like James’. Along similar lines, if there are many “knowledge concepts,” none of which enjoys any special theoretical advantage, the task of articulating the contours of *one* of them (e.g., that of rich, educated Westerners) seems to be at best of anthropological interest.

²³Whether this skeptical challenge can be met head on is, unsurprisingly, a matter of controversy. Alexander et al. (Forthcoming) argue that it cannot. Williamson (2011) argues that it can.

²⁴See Sosa (2007), though I don’t know if Sosa would endorse the particular version of the position I consider here.

One way of bringing this point out involves showing how demographic variation in intuitions about knowledge undermines an otherwise plausible defense of a certain sort of philosophical methodology—one originally associated with ordinary language philosophers, but versions of which are still quite common today. The defense I have in mind was offered by J.L. Austin:

[O]ur common stock of words embodies all the distinctions men have found worth drawing, and the connexions they have found worth marking, in the lifetimes of many generations: these surely are likely to be more numerous, more sound, since they have stood up to the long test of the survival of the fittest, and more subtle, at least in all ordinary and reasonably practical matters, than any that you or I are likely to think up in our arm-chairs of an afternoon – the most favoured alternative method.” (Austin, 1956)

When we find out, however, that different demographic groups have found slightly different distinctions worth drawing, it’s quite natural to worry that there’s nothing especially “sound” or “subtle” about the distinctions that *we* are inclined to draw, at least as compared to those drawn by other demographic groups. In light of this worry, pressing on and trying to systematize *our* intuitions would be hasty; we would run the risk of engaging in a largely worthless task.

In the next section, I’ll offer a strategy for showing that epistemological questions are substantive. I hope to establish that a great many epistemological questions are immune from the sort of challenges I’ve been discussing so far. However, the strategy will not work in all cases. In particular, it will not apply to the particular examples I’ve discussed so far; for all I say in the next section, debates between internalists and externalists and debates over the proper epistemological classification of cases like those discussed by Weinberg et al. may be merely verbal. But the lesson to draw from this—I’ll argue—is that these debates are in fact much less typical and much less central to the epistemological enterprise than they might at first seem.

3.3 Vindicating Epistemological Debates

In this section I’ll pursue a two-part strategy for defending epistemological questions from the charge of mere verballity. The first part will involve offering a coherence constraint that applies to our views about which questions are merely verbal, and which are substantive. The constraint is quite modest—in particular, it is powerless to resolve disputes in metametaphysics. Nevertheless, the second part of my strategy will involve arguing that if the constraint is appropriate, then we are committed to regarding a great many epistemological questions—including questions about justification and knowledge—as substantive. I’ll go on to explain why the debates discussed so far—debates over the proper classification of examples like Claire’s, Hal’s, and those discussed by Weinberg et al.—are atypical, and can’t be vindicated using the strategy I defended.

3.3.1 Transmission of Substantiveness

Suppose we come upon two people engaged in a debate about whether a particular country's government is legitimate. If we don't take it to be immediately obvious whether the question they are disagreeing about is substantive or merely verbal, how might we try to determine its status? A natural strategy is to ask the participants to the debate what *further* questions they think turn on the question of whether or not the government is legitimate. They might say that the question of whether the government is legitimate is relevant (at least *ceteris paribus*) to questions such as whether the government in question should receive foreign aid, whether the citizens of the country have an obligation to follow its laws, whether a humanitarian military intervention in that country might be just, and other ground-level moral questions. These latter questions, we might think, are clearly substantive—taking a stand on them amounts to taking a stand on matters of practical significance.²⁵ If we take these latter questions to be substantive, and we take the question of whether the government is legitimate to bear on them, then it seems that we must—on pain of incoherence—take the question of whether the government is legitimate to be itself substantive.

More generally, it's not clear what we'd be thinking if we thought that some particular question was not a substantive one, while also thinking that resolving that question would have important consequences for other questions that *are* substantive. If merely verbal questions are distinctive in that nothing of substance turns on such questions, then we must think that a question is not merely verbal if we think that it is relevant to some *other* question that *is* substantive.

The following principle, which I'll call '(TS)', for *transmission of substantiveness*, is intended to capture the above line of thought:

(TS) If you think the question of whether *Q* is a substantive one, and you take the question of whether *P* to be relevant to the question of whether *Q*, then you are rationally required to think the question of whether *P* is substantive one as well.

It's not immediately clear how to understand (TS)—in particular, we might ask how to understand the phrases 'relevant' and 'rationally required.' I'll address these questions in turn. First, however, I want to say a bit about why I've framed (TS) as a coherence constraint concerning which questions we must *think* are substantive (given facts about which other questions we *take them* to be relevant to), rather than as a constraint on which questions must *be* substantive (given facts about which other questions they *are* relevant to). The short answer is that I view the notion of someone's *taking* a question to be relevant to another one as a more proper starting point for analysis than the notion of a question's *being* relevant to another one. While the following biconditional is extremely plausible, I'm inclined to view the right-to-left direction as the more explanatory one:

²⁵I assume that James' pragmatic criterion for what it takes for a question to be substantive is at least a sufficient one, if not a necessary one.

- Question Q is relevant to question $P \leftrightarrow$ Question Q is rightly taken to be relevant to whether P

Here's a first pass characterization of what I have in mind by 'relevant'—one takes the question of whether P to be relevant to the question of whether Q just in case one is disposed to change one's mind about whether Q upon changing one's mind about whether P . This doesn't capture quite what I have in mind, for a reason that the following example brings out. Suppose Walter has a stable disposition to engage in wishful thinking. Walter is stably disposed to become more confident that it *is* the case that P when he becomes more confident that it would be *good* if it were the case that P . But this isn't because he engages in anything like a process of reasoning. It's not that he has a standing belief that things tend to be for the best; his beliefs don't license reasoning from the premise that it would be best if P to the conclusion that P , as would the beliefs of Dr. Pangloss. Rather, in Walter's case the transition is an unconscious, non-deliberative one—we may stipulate that if he noticed the process operating in any particular instance, his better judgment would kick in and override his wishful thinking.

In the sense of 'relevant' that I intend in (TS), Walter does not count as taking the question of whether it would be good that P as relevant to the question of whether P just because he's disposed to engage in wishful thinking. While Walter is disposed to change his mind about whether P upon changing his mind about whether it would be good that P , the disposition isn't of the right sort.²⁶ For a disposition to change one's mind about Q upon changing one's mind about P to constitute one's regarding P as relevant to Q in the right sense, it must be a disposition that could be manifested in a process of reasoning. We might say that it's only if one takes P to be a *reason* to become more (or less) confident that Q that one counts as regarding P as relevant to Q —where the notion of taking something to be a reason is one that can be cashed out psychologically, in terms of one's being disposed to change one's mind in certain distinctive ways (i.e., not via wishful thinking or some other deviant route, but via reasoning).

Now that we've said a bit more about 'relevant,' we can turn to 'rationally required.' The intuitive idea is that having attitudes that fail to conform to (TS) amounts to a kind of inconsistency, or incoherence. But I don't think it's important for my purposes that I say too much about just what kind of incoherence is at issue; my hope is that I can use (TS) to do some work without entering into debates about the normativity of rationality.²⁷ Ultimately, my aim is to use (TS) to convince you, my reader, to regard certain epistemological questions as substantive. All that's required for (TS) to serve this purpose is that you be disposed, upon realizing that you have attitudes that (TS) says you're rationally required not to have, to come to modify your attitudes so that they conform to (TS).²⁸ As long as you have these

²⁶This is a special case of the "deviant causal chains" problem, perhaps best known from the work of Donald Davidson. See Davidson (1980).

²⁷I have in mind the debate between Broome (1999) and Kolodny (2005).

²⁸This is a bit crude—it will matter *how* you modify your attitudes so that they conform to (TS). But as long as you hold fixed certain commitments that I expect you'll want to hold fixed, (TS) will

dispositions (and I suspect that you do), it doesn't for my purposes matter exactly what sort of fault is involved in having attitudes that one is rationally required not to have.

(TS) does not provide anything like a general recipe for resolving debates about which questions are merely verbal, and which are substantive. To see why not, consider the following example. Suppose Rudolf and Ted disagree about whether the debate over P is substantive, or merely verbal; Rudolf thinks it is merely verbal, while Ted thinks it is substantive. Suppose that both parties accept (TS), and both parties regard the question of whether P as relevant to the question of whether Q . Ted takes this fact about relevance to support the position that the question of whether P is substantive, since Ted thinks that the question of whether Q is substantive. But one man's modus ponens is another's modus tolens—Rudolf thinks that the question of whether Q is merely verbal, and therefore doesn't take the connection between P and Q to indicate that either question is a substantive one.

As long as disagreements about which debates are substantive run deep enough, we cannot resolve these disagreements by appeal to (TS). It's only when two parties both agree that some particular question is substantive, and agree that certain other questions are relevant to it, that they can appeal to (TS) to generate further shared commitments. In the cases about which metametaphysicians tend to disagree, this sort of common ground is typically absent. For instance, those who regard the debate between mereological nihilists and mereological universalists as a substantive one might point out that it has consequences for questions such as how many objects the world could have contained, which might seem like a substantive question. But those initially inclined to regard the debate between nihilists and universalists as merely verbal will probably, upon appreciating its connection to questions about how many objects the world could have contained, come to see these latter questions as merely verbal as well.²⁹

The fact that (TS) is impotent to resolve vexed metametaphysical questions is, I think, a point in its favor—in endorsing it, I don't thereby commit myself one way or the other on any standing controversies. Nevertheless, despite the innocuousness of (TS), I'll argue that it can do some significant metaepistemological work. The basic strategy is straightforward—I'll argue that we treat many epistemological questions as relevant to uncontroversially substantive non-epistemological questions. By (TS), then, we're committed to regarding the epistemological questions themselves as substantive.

3.3.2 Applying (TS) to Epistemological Questions

A great many epistemological debates are (or can be naturally recast as) debates about whether the belief that P has some positive epistemic status (rationality, jus-

require you to regard many epistemological debates as substantive.

²⁹Here we might recall the words of David Lewis: "any competent philosopher who does not understand something will take care not to understand anything else whereby it might be explained." (1986, p. 203, fn. 5)

tification, evidential support), for some claim P that itself is uncontroversially substantive.³⁰

For instance, classic epistemological questions include questions about whether we have any good reason to believe that there exists an external world, (Descartes, 2004) or that bread will continue to nourish us (Hume, 2000). The questions of whether there is in fact an external world, or whether bread will in fact continue to nourish us, are uncontroversially substantive. Perhaps their answers are obvious, but they don't strike us as questions that are empty, or somehow terminological.

In recent years questions about how we ought to respond to disagreement have been receiving a great deal of attention in the epistemological literature.³¹ These questions are naturally cast in the form mentioned above; we can put these questions as questions whether is it rational for one to suspend judgment concerning whether P when one is faced with disagreement from one's epistemic peers concerning whether P , and where the proposition that P concerns some straightforwardly substantive matter.

Many epistemological debates in the philosophy of science also have this feature; they are often debates about whether our evidence warrants belief in some claim P , where the question of whether P is itself (almost) uncontroversially substantive. Unless we're verificationists, we'll think that the question of whether electrons exist or the world is merely observationally *as if* electrons exist, is a substantive question. So debates between constructive empiricists and scientific realists about whether we ought to believe that electrons exist will be debates about whether we ought to believe that P , for some substantive claim P .³² Along similar lines, we might consider

³⁰ *Terminological note:* I'll alternate freely between the locutions 'it is rational for one to believe that P ,' 'one ought to believe that P ,' 'it is reasonable for one to believe that P ,' 'one has justification to believe that P ,' and 'the rational attitude given one's evidence is to believe that P .' I don't deny that these locutions can fruitfully be given distinct senses, but I do think that this is a matter of stipulation. I don't think ordinary language already distinguishes them. I suspect that Alston, at least, would have no objection here. After surveying a list of quotations from various epistemologists, he writes: "Some of these formulations are in terms of 'rational' rather than 'justified', but I believe that a close reading of the context in each case will reveal that the authors who use 'rational' mean to be marking out a property of beliefs that has the same, or a very similar, epistemic force as what the other authors are trying to get at with 'justified'." (2005, p. 12) It is true that for some purposes, it is very important to distinguish between "it is rational for one to believe that P " and "one ought to believe that P "—in particular, in debates about whether rationality is "permissive," in the sense discussed by White (2005). And this might seem quite important to the arguments below as well, since permissivists will not generally engage in inferences from "it is rational for me to believe that P " to " P "—if they ever find themselves in cases they take to be permissive, then such an inference pattern would lead them to contradiction. But I don't think this should create much confusion—permissivists can still engage in the inference pattern from "it is *uniquely* rational for me to believe that P " to " P ", and the availability of such an inference is enough for my arguments in the rest of this section to get traction.

³¹ See the essays collected in Feldman and Warfield (2010).

³² I have in mind the debate between van Fraassen (1980) and his critics. Of course, to the extent that we're skeptical about Van Fraassen's version of the belief/acceptance distinction, we may be skeptical about the substantiveness of questions about whether to believe (as opposed to merely accept) various claims about unobservables. Melchert (1985), Horwich (1991), and others have expressed skepticism along these lines.

epistemological debates between Bayesians and Frequentists about the appropriate methodology in hypothesis testing. If we take their debate at face value, the Bayesian and the Frequentist seem to disagree over which beliefs it would be rational to adopt in response to certain bodies of evidence. In extreme cases, the Frequentist might recommend that we become more confident that some coin has a certain bias—clearly a substantive matter—while the Bayesian would give the opposite recommendation.³³

Even many epistemological debates that might seem *recherché* can be cast in this form. For instance, the Sleeping Beauty problem concerns which beliefs about whether a coin landed heads—clearly a substantive question—it would be rational to have in some odd (but not *too* far-fetched) circumstances (Elga, 2000).

Why does it matter, for my purposes, that so many epistemological debates have this form? In each of these cases, we can appeal to (TS) to argue that we're committed to regarding the epistemological debate in question as substantive. This is because most of us take the question of whether we ought to believe that *P* as relevant to the question of whether *P*. Myles Burnyeat makes this point in the context of a discussion of skepticism:

To take a well-worn, traditional example, if the evidence of our senses is really shown to be unreliable and the inferences we ordinarily base on this evidence are unwarranted, the correct moral to draw is not merely that we should not claim to know things on these grounds but that we should not believe them either. Further, in the normal case, that which we think we should not believe we do not believe...If skepticism is convincing, we ought to be convinced, and that ought to have a radical effect on the structure of our thought.³⁴ (Burnyeat, 1983, pp.118-119)

In my terminology, Burnyeat is suggesting that at least in the normal case, one takes the question of whether one ought to believe that one's senses are reliable as relevant to the question of whether one's senses *are* reliable. In the normal case, upon concluding that one ought(oughtn't) believe that one's senses are reliable, one will reason one's way to the conclusion that they are(aren't) in fact reliable, and will continue(cease) relying on them.

What Burnyeat says about skepticism is just as plausible when applied to the other examples of epistemological debates mentioned above. I suspect part of the reason that the disagreement debate has attracted so much attention is that the questions it poses are immediately relevant to straightforwardly substantive non-epistemological questions. If we come to think that beliefs are unjustified when they are held in

³³See Howson and Urbach (1996). The example I have in mind is discussed in the section entitled 'A Well-Supported Hypothesis Rejected in a Significance Test.'

³⁴I suspect that Burnyeat is not using "reliable" in quite the way that most epistemologists use it today. In particular, because skeptical arguments typically don't aim to show that our senses *in fact* typically lead us astray (only that we have no reason to rule out the hypothesis that they do so), I suspect that Burnyeat is using "reliable" to mean something more like "appropriate to rely on" rather than "typically delivering accurate results." If we read him this way, we can make sense of the idea that the skeptic aims to show that our senses are unreliable (and not merely to show that we can't know whether they're reliable or not).

the face of certain sorts of disagreement, we will likely go on to suspend judgment concerning a wide range of controversial political, economic, and scientific matters.³⁵

Along the same lines, at least normally, (and I'm not sure whether there are any *abnormal* cases) we'll treat questions of whether we *ought* to believe various claims about unobservables as relevant to questions about whether those claims are *true*. Even in the case of Sleeping Beauty, if we've settled on an answer to the problem and we later find ourselves in Beauty's predicament, our views about how likely the coin is to have landed heads will likely be informed by our epistemological views about what it would be rational to believe in such circumstances. It's true that in the case of the Sleeping Beauty problem, our views about non-epistemological, clearly substantive matters—which way some coin landed—wouldn't be affected until we found ourselves in Beauty's predicament.³⁶ But it's very much in the spirit of (TS) to hold that if, in coming to an epistemological conclusion, one is now disposed upon finding oneself in certain circumstances to reason one's way to a new substantive non-epistemological conclusion, then one must regard the original epistemological question as substantive even if one isn't already in the relevant circumstances.

In each of these cases of epistemological questions about whether it is rational to believe some claim *P*, if we hold fixed that (1) we believe that the question of whether *P* is a substantive one, and (2) we take the epistemological question to be relevant (in the sense spelled out in the previous section) to question of whether *P*, then the only way for us to satisfy (TS) is to regard the epistemological question as substantive as well. Moreover, in each of the examples I mentioned, it's quite plausible that we *should* hold (1) and (2) fixed—that is, we're right to regard the first-order, non-epistemological questions as substantive, and we're also right to take the epistemological questions to be relevant to those first-order, non-epistemological questions.

Granted, the idea that we're right to treat epistemological questions as relevant to non-epistemological questions is *itself* easier to explain on certain epistemological views than others. On many epistemological views, one might have justification to believe that one has justification to believe that *P*, while one in fact fails to have justification to believe that *P*.³⁷ On such views, in reasoning from the premise that one has justification to believe that *P*, to the conclusion that *P*, one might be led from a justified true belief in the premise, to an unjustified false belief in the conclusion.³⁸ But I take it that any epistemological view—if it is to be plausible—must offer an explanation of why it is generally reasonable to take epistemological

³⁵In fact, this is not completely straightforward. See Elga (2007) on the charge of “spinelessness.” The general point still stands, however—questions about the epistemic significance of disagreement are typically relevant to questions about non-epistemological matters.

³⁶Though we might say that even prior to finding ourselves in Beauty's predicament, forming a view about the rational credences to have in such a predicament would influence our plans about what to believe in that predicament. Here I'm appealing to the notion of a ‘plan’ for belief in the sense explicated by Gibbard (2003).

³⁷This is true especially, though not exclusively, on broadly externalist views.

³⁸Views that endorse “level-bridging” principles, such as the principle that if one ought to believe that *P*, then one ought to believe that one ought to believe that *P*, and/or vice versa, have an easier time explaining why the sort of reasoning I've been discussing is good reasoning.

questions about one's justification/evidence for P to bear on the question of whether P ; if epistemological questions were not generally relevant to first-order questions in this way, it would be much harder to see why we should care about them. To the extent that certain views can't explain this relevance, that constitutes an objection to such views.

I've argued that we must think that a great many epistemological questions about what it is rational to believe are substantive, if we accept (TS) and some other plausible assumptions. We might wonder, then, whether (TS) might be used to vindicate the epistemological debates discussed earlier—debates between internalists and externalists about justification, and debates over whether subjects in Gettier, truetemp, and cleverly disguised mule cases have knowledge or mere belief. I'll explain in the next two subsections why (TS) does *not* straightforwardly commit us to regarding these debates as substantive.

3.3.3 Internalism and Externalism

Recall the examples of clairvoyant Claire and Hallucinating Hal discussed earlier. Internalists and externalists disagree over whether the subjects of examples like these have justified beliefs—internalists think that Claire is unjustified and Hal is justified, while externalists hold the opposite set of views. But could this difference ever manifest itself in different dispositions to reason one's way to non-epistemological beliefs? It's far from obvious, as the following discussion will elucidate.

Suppose our internalist finds herself in a situation like Claire's. While the internalist will believe that the president is in New York, if she reflects on her epistemic situation and deliberates with the aim of holding a justified belief on the matter of the president's whereabouts, she'll likely abandon this belief; by her lights, the belief will seem to be unjustified. However, the same is true for the externalist. Suppose our externalist finds himself in a situation like Claire's—he believes that the president is in New York, and in fact this belief was caused by a reliable process, but he has no reason to suspect this. If the externalist reflects on the epistemic status of his belief he'll be disposed to abandon it; by his lights, the belief is no more likely to be right than a guess, since he doesn't take it to have been caused by a reliable process. The upshot is that the internalist and the externalist would reason their way to the same attitudes concerning the president's whereabouts—each would abandon the belief that the president is in New York—were they to find themselves in a situation like Claire's.

What if they found themselves in a situation like Hal's? In such a case, the internalist would maintain her belief that there is an apple on the table—by her lights, this belief would be justified, since supported by excellent perceptual evidence. The externalist would also take his belief to be justified; he would take it to have been caused by visual perception, a reliable process. As in the earlier case, both the internalist and the externalist would, if they reflected on their epistemic situation with the aim of retaining only justified beliefs, come to the same conclusions.

To be clear, in this section I've been assuming that the externalist can meet the challenge I mentioned at the end of the previous section; I assume that the externalist

can find a way to vindicate the practice of reasoning from premises about the epistemic status of one's belief that *P*, to conclusions concerning whether *P*. To the extent that the externalist *cannot* vindicate such practices, perhaps she won't reason in the way I've been discussing in these cases. But this would be a problem for externalism, as being able to engage in such practices is arguably a central purpose of forming epistemological judgments in the first place. While I will continue to assume that the externalist can meet the challenge, this is not uncontroversial—Karl Schafer (Ms.), for one, has argued that the externalist *cannot* meet a similar challenge, and that this failure constitutes a serious drawback of externalism.

I take it that the above considerations make it plausible that typical internalists and typical externalists won't reason their way to different non-epistemological conclusions upon finding themselves in cases like Claire's and Hal's (modulo the assumption just mentioned). I don't pretend to have shown that disagreements over how to classify cases like Claire's and Hal's will *never* manifest themselves in non-epistemological disagreements. I do, however, hope I've made it clear that there is a big difference between the internalism/externalism debate on the one hand, and the debates discussed in the previous section on the other. Our views on the debates in the previous section can *clearly* manifest themselves in our dispositions to form uncontroversially substantive non-epistemological beliefs. But if debates over how to classify examples like Claire's and Hal's can ever manifest themselves in our dispositions to reason our way to non-epistemological beliefs—which I doubt—the path is much less straightforward.

Of course even if I'm right, these observations fall far short of an argument to the effect that such disagreements are merely verbal—all I've argued is that a strategy that works to show that *many* epistemological questions are substantive fails in this case. But other strategies might work in this case—for example, we might hold that the internalism/externalism debate is substantive because internalists and externalists are committed to holding different epistemic reactive attitudes towards people like Claire and Hal; perhaps the internalist's position commits her to blaming Claire and admiring Hal, while the externalist's position commits her to the reverse. While I'm skeptical of views that tie epistemological judgments to reactive attitudes, nothing I've said so far tells against this strategy.

Still, my skepticism isn't unusual. Alston (1988) argues that the applicability of deontological concepts—a class in which he includes concepts like blameworthiness and praiseworthiness—in epistemological contexts would require that we have voluntary control over our beliefs, which he argues that we lack. While conceding that we generally lack voluntary control over our beliefs, Feldman (2001, 2008) defends the applicability of *some* deontological concepts in the context of epistemology—he holds that it's sometimes true to say that we ought, or have justification, to believe various claims—while expressing skepticism about links between these concepts on the one hand, and blame and praise on the other.

Even if Alston and Feldman are wrong, and blame and praise *do* have important epistemological roles, the observations of the present subsection would point the way to some promising avenues for future epistemological research. If certain debates between internalists and externalists can only be shown to be straightforwardly

substantive if we can establish a tight connection between epistemological judgment on the one hand, and blame and/or praise on the other, that would be an interesting and important conclusion in its own right, even if the project ultimately *can* be successfully carried out. Once we take seriously the idea of blame and praise in an epistemological context, a number of questions arise. How do epistemic blameworthiness and praiseworthiness differ (if at all) from the phenomena of blameworthiness and praiseworthiness that figure in moral theorizing? Is the world a better place when rational people prosper, or when the foolish suffer? What roles should epistemic blame and praise play in regulating social interactions? If we hold that certain epistemological questions are substantive because of their links to blame and praise, questions like the ones above should seem both interesting and pressing.

3.3.4 Knowledge

In §3.2 I argued that a wide range of debates about what it is rational (justified, warranted) to believe can be defended from the charge of mere verballity, because such debates are relevant to straightforwardly substantive, non-epistemological questions. Can a similar strategy be used to defend the status of debates about which beliefs constitute knowledge?

There's a tradition in epistemology that sees the utility of the concept of knowledge as issuing primarily from its role in enabling us to identify reliable informants. While this tradition perhaps reached its apotheosis in Edward Craig's *Knowledge and the State of Nature* (1990), it has also been defended by Bernard Williams (1970), Robert Brandom (2000, chap. 3), and Allan Gibbard (2003, chap. 11).³⁹

According to the line of thought pursued by Williams, Craig, Brandom, and Gibbard, a (perhaps *the*) central role of the concept of knowledge is in helping us form beliefs about non-epistemological matters. In trying to determine whether P , we often seek the guidance of people who know whether P . Determining who knows whether P is not an idle classificatory exercise, irrelevant to non-epistemological matters. Rather, it is a crucial step in our worldly inquiries; if we determine that someone knows whether P , then we'll generally come to believe that P if she asserts that P , and to believe that $\sim P$ if she asserts that $\sim P$. If we find this way of thinking about knowledge attractive, then we'll generally take questions about who knows what to be substantive; much the same strategy that I used in §3.2 to argue that debates about rationality/justification/warrant are substantive will apply to debates about knowledge as well. However, this strategy will not extend to a certain class of questions, of which the questions studied by Weinberg et al. are members. These are the questions concerning how to sort people into knowers and mere believers when we occupy what Williams calls the "examiner situation":

Academic writings about knowledge are notably fond of that which might be called the *examiner* situation: the situation in which I know that P is true, this other man has asserted that P is true, and I ask the

³⁹And likely a good many others.

question whether this other man really knows it, or merely believes it. I am represented as checking on someone else's credentials for something about which I know already...But this is far from our standard situation with regard to knowledge; our standard situation with regard to knowledge (in relation to other persons) is rather than of trying to find somebody who knows what we don't know; that is, trying to find somebody who is a source of reliable information about something...Our standard question is not "Does Jones know that *P*?" Our standard question is rather "Who knows whether *P*?" (Williams, 1970, p. 146)

Williams doesn't tie his discussion of the oddity of the examiner situation to worries about mere verballity—the passage above appears in a paper primarily concerned with doxastic voluntarism. However, the strategy I've been pursuing so far suggests a reason for skepticism about the substantiveness of certain questions about knowledge asked from the examiner situation. When we occupy the examiner situation, a central reason for asking whether somebody knows that *P* is absent, since we have no need to rely on testimony concerning whether *P*. Correlatively, a central reason for regarding most questions about knowledge as substantive—the fact that deciding them one way or the other will make a difference to our dispositions to form non-epistemological beliefs—doesn't apply to questions asked from the examiner situation.

Of course, determining whether to rely on someone's testimony isn't the *only* reason to ask who knows what. If I am engaging in illicit activities, I might be interested in who knows what I've been up to. This isn't because I want to rely on their testimony (*I* know what I've been doing), but because I want to make sure that my secret doesn't get out—if people know what I've been doing, they make speak up, and they're more likely to convince other people of my wrongdoing if they have knowledge rather than mere suspicion.⁴⁰ Or, even if I know whether *P* and so don't need to rely on testimony concerning whether *P*, I might be interested in who else knows whether *P* because I *do* need to rely on testimony concerning *other* questions in the same subject matter as *P*. For instance, I might be interested in who knows who plays first base for the Yankees, not because I need to rely on testimony concerning *this* fact, but because I want to find somebody who's knowledgeable about baseball more generally, so I can rely on his or her testimony about *other* baseball-related matters. But many of the "examiner situation" cases that epistemologists focus on lack these features. Recall the Gettier case described by Weinberg, Nichols, and Stich. When we ask whether Bob knows or merely believes that Jill drives an American car, it doesn't seem as if different answers to our question are supposed to correspond to different predictions about Bob's future behavior, or his ability to influence the opinions of others, or his reliability concerning other automotive matters.

The relevance of the present discussion to the experimental critique should now be starting to come into focus. Research into variance in epistemological intuitions across demographic lines has—at least as far as I know—exclusively involved the examiner situation; experimenters describe a situation in which it is stipulated that

⁴⁰See Williamson (2000, p. 62) for a related discussion.

P, and subjects are asked whether some character in the situation knows that *P* or merely believes it. In these sort of situations, questions about who knows what lack their typical, clear connections to non-epistemological questions. As a result, it's far from obvious whether differences between—for example—Westerners and East Asians on these epistemological questions will manifest themselves in different dispositions to rely on testimony in forming non-epistemological beliefs, or different predictions about how people will behave. There is, then, some pressure to suspect that the disagreement between a typical Westerner and a typical East Asian in one of Weinberg et al.'s studies is merely verbal.

Moreover, even if there *weren't* any such disagreement—even if, for example, Nagel (Forthcoming) is right that there are no stable demographic dispositions in epistemological intuitions in cases of the sort studied by Weinberg et al.—there would *still* be reason to worry about the substantiveness of questions about knowledge asked from the examiner situation. Ultimately, while actual demographic differences may make salient the possibility that questions asked from the examiner situation are merely verbal, the reasons to worry that they *are* merely verbal—namely, that they lack inferential connections to non-epistemological questions—apply whether or not people actually disagree about such questions.

Once the distinction is made between questions asked from the examiner situation and other questions about knowledge, we may see the position according to which there is a multitude of “knowledge concepts” in a new light. In §2.2, we encountered the threat that admitting that there are various different concepts of knowledge—akin to various different dialects of English—would lead to the result that epistemological questions about who knows what are often merely verbal. But we now have a principled way of containing the threat. Many questions asked from the examiner situation may well turn out to be merely verbal—as already mentioned, these questions lack straightforward connections to non-epistemological questions. But they are unusual in this respect. Most questions about knowledge must be substantive, since they are relevant to our decisions about whom to trust, and the (obviously substantive) testimony-based beliefs we go on to form as a result of these decisions. If it turns out that there is significant disagreement across demographic lines even on *these* epistemological questions—a claim for which there is as yet no empirical evidence—then there will be difficult questions about how best to respond to that disagreement. But that would be a very different situation from the one we find ourselves in now; as matters stand, the only epistemological disagreements that seem to track demographic lines are ones that we can regard as merely verbal without thereby committing ourselves to regarding *most* epistemological disagreements as merely verbal.⁴¹

⁴¹In some respects, my strategy in this section has been similar to that of Mark Kaplan in “It’s Not What You Know that Counts” (1985). In that paper, Kaplan argues that questions about knowledge (as opposed to questions about justification) are unimportant and uninteresting, largely because they are irrelevant to questions about how to go about conducting our wordly inquiries. While I am sympathetic to the idea that epistemological questions *should* be relevant to questions about how to inquire, *and* to the idea that a large class of questions about knowledge that epistemologists have focused on fail to meet that standard of relevance (namely, the class of questions asked from the examiner situation), I cannot endorse all of Kaplan’s conclusions, in large part because it seems to

My strategy in this chapter so far has been one of damage containment. Even if we concede that debates between epistemological internalists and externalists are merely verbal, and that the epistemological questions focused on by experimental philosophers are merely verbal, there are principled reasons for denying that the threat of mere verballity in epistemology generalizes. Epistemological questions are uncontroversially substantive when they are relevant to other questions that are uncontroversially substantive, and I've argued a wide range of epistemological questions *are* relevant to uncontroversially substantive questions of fact. Whether the challenges raised by Alston and the experimentalists can be met is an interesting question—but the fate of epistemology does not rest on answering it.

It may seem as if I've appealed to unnecessarily limited materials in defending various epistemological debates as substantive. Even if we grant that my defense succeeds, it may seem as if I've been fighting with one hand behind my back; after all, it's not as if epistemological questions are *only* substantive when they bear on non-epistemological questions, is it?

In the final section of this chapter I'll sketch a positive metaepistemological view according to which, in fact, it *is* the case that epistemological questions are only substantive when they are relevant to non-epistemological questions. My claims in the next section will be independent of what I've argued so far—one can accept my defense of various epistemological questions as substantive without accepting the claim that *only* the questions to which my defense applies are substantive. Nevertheless, I think it will be fruitful to see how a certain sort of metaepistemological view naturally leads to skepticism about the substantiveness of epistemological questions that are not relevant to non-epistemological questions.

3.4 Epistemic Expressivism

In this chapter I've stressed the idea that we typically take epistemological questions as relevant to non-epistemological questions. From the deliberative point of view, concluding that my evidence overwhelmingly supports the claim that *P* settles the question of whether *P*. If someone asserted that her evidence overwhelmingly supported the claim that Obama will be reelected, but went on to say that she thought that her evidence was misleading and that in fact Romney will run and win, it would be natural to suspect her of confusion or incoherence, rather than of merely taking too seriously an implausible hypothesis. This isn't uncontroversial—some philosophers think there's nothing confused or incoherent about believing claims of the form "*P*, but my evidence on balance supports the claim that $\sim P$."⁴² For present purposes, however, I will take as a datum that there *is* something confused or incoherent about believing such claims. Once we do so, some interesting metaepistemological consequences follow.

me that he fails to appreciate that we do not *always* occupy the examiner situation when we ask who knows what.

⁴²See, e.g., Weatherson (2008).

In particular, it is hard to explain this datum if we adopt the following metaepistemological picture, which I take to be widely held. According to this picture, epistemology is helpfully thought of as concerning what it takes for beliefs to have certain “epistemic properties,” such as the property of being justified, or being supported by the evidence. Moreover, for most propositions P , questions about which epistemic properties are exemplified by someone’s belief that P are logically independent of the claim that P ; sometimes P is true but the evidence doesn’t support it, and vice versa. If this is our metaepistemological picture, then it is hard to see why there should necessarily be anything odd about believing that P , while believing that one’s evidence on balance supports the claim that $\sim P$. In general, when two claims are logically independent, then even if it is *normally* reasonable to believe one only if one believes the other, there need be no confusion or incoherence involved in believing one without believing the other, and there will be *some* situations (perhaps unusual ones) in which that is the reasonable combination of beliefs to hold. For instance, “where there’s smoke, there’s fire” is a good rule of thumb. But the claim that there’s smoke in some area is logically independent from the claim that there’s fire near that area, and in some situations we should believe that there is smoke but no fire—when in the presence of someone operating a bee smoker, for instance.

The rule “when one’s evidence supports the claim that P , P ” doesn’t seem to be like this. There don’t seem to be situations in which it’s reasonable to believe that one’s evidence supports the claim that P , but not to believe that P . We might account for this by positing an epistemic level-bridging principle—e.g., the principle that if the evidence supports the claim that P , then the evidence *also* supports the claim that the evidence supports the claim that P —as a brute epistemological fact. But this wouldn’t be very satisfying—for most logically independent claims P and Q , sometimes it’s reasonable to believe P but not Q , and adopting level-bridging principles as brute gives us no explanation for why this generalization fails to apply when Q concerns the character of one’s evidence for P .⁴³

A more attractive response, I think, is to reject the metaepistemological picture that leads to the puzzle; we should reject the idea that epistemology is most helpfully thought of as involving questions about the distribution of epistemic properties, where facts about which beliefs have those properties are typically logically independent of the contents of those beliefs.⁴⁴ What should we put in its place? While offering a detailed alternative is well beyond the scope of this dissertation, it’s possible to quickly give a flavor of the position I prefer.

On the view I favor, epistemology is better thought of as a tool for helping us structure our non-epistemological inquiries than as a domain with its own distinctive subject matter.⁴⁵ Prior to introducing epistemological terms into our language (or

⁴³To be clear, I do agree that the level-bridging principle in question is true. But I also think that we can explain *why* it must be true. I tried to go some way towards providing such an explanation in the previous chapter, and I continue that effort in the remainder of the present chapter.

⁴⁴This isn’t to say that we shouldn’t endorse a “quasi-realist” style metaepistemological picture, in which ultimately (after a good deal of explanatory work), we “earn the right” to talk in terms of epistemic properties. See Blackburn (1984).

⁴⁵This idea can be fleshed out in many ways, most of which will fall under the label of “expres-

epistemological concepts into our conceptual repertoire), we can treat some inferences as good and others as bad, but we lack certain helpful resources for talking about (or thinking about) this practice in general. For instance, I might be disposed to infer that Zelda is a duck from the claims that Zelda looks like a duck, swims like a duck, and quacks like a duck, and that Helga is a duck from the claims that Helga looks like a duck, swims like a duck, and quacks like a duck. But if I lack epistemological terms, I will be unable to explicitly identify these inferences as having something in common—I won't be able to say, for instance, that in general, something's exhibiting duckish behavior and manifesting a duckish appearance is *evidence* that it is a duck.

Once we have epistemological terms, we can *explicitly* endorse (or reason about whether to endorse) certain inferences that we could before only accept *implicitly* (i.e., by making them). On this picture, the purpose of epistemological notions isn't to let us talk about some distinctive subject matter that we'd be unable to refer to without such notions (e.g., the distribution of epistemic properties)—rather, they are for helping us explicitly talk and reason about which inferences concerning *non-epistemological* matters to accept. I said earlier that the metaepistemological picture I favored has consequences for the question of which epistemological questions are substantive. We're now in a position to see why this is.

The following analogy should help.⁴⁶ In elementary chemistry textbooks, atoms are described as having valences, or valence numbers, which measure the number of bonds they can form. For example, an atom with a valence of 1 is missing an electron, and can form a bond with another atom with a valence of -1 , since the latter would have an electron to spare. The entire theoretical role of valence numbers is to explain bonding behavior along the above lines. Once we understand this, we'll realize that 'alternative' theories of valences on which, e.g., all the signs are flipped, but the predicted bonding behavior is just the same, are in fact not substantive alternatives at all.⁴⁷ Because valences are just tools for predicting bonding behavior, there cannot be 'competing' theories of valences that always agree about bonding behavior.

But on the expressivist metaepistemological picture I've been sketching, epistemological notions are a tool for helping us deliberate about non-epistemological matters in much the same way that valence numbers are a tool for helping us predict bonding behavior. If we accept this, then we should also think that 'competing' epistemological theories with the feature that accepting one rather than the other would never lead to our reasoning our way to different non-epistemological beliefs are in fact not in any competition at all. The choice between such theories should be made on grounds of notational convenience—it shouldn't be thought of as a substantive one.

sivism." The following sketch draws heavily on the work of Robert Brandom (1994, 2000). See also Blackburn (1984) and Gibbard (1990, 2003).

⁴⁶The use of this analogy is inspired by Dennett (1987b).

⁴⁷Note that valence numbers are plausibly, in this respect, quite different from unobservable entities that are hypothesized to cause observable behavior. Unless we're verificationists, we'll probably agree that there are often substantive disagreements between theories that disagree about the behavior of unobservable entities, but agree about the behavior of observables. If this our view, we won't think that unobservable entities should be thought of as mere tools or posits we use only to ease prediction of observable goings-on.

If the arguments of §3.1-2 are sound, adopting the expressivist metaepistemological picture I've sketched would do little violence to first-order epistemology as currently practiced—we could go on in good conscience having the most of the same epistemological debates we were already having, with the added benefit of having a better understanding of how such debates fit into the rest of our cognitive lives. However, adopting the metaepistemological picture I've been exploring might do *some* violence to contemporary epistemological practice—some epistemological debates may be inferentially insulated from non-epistemological questions. I hope my discussion of certain aspects of the internalism/externalism debate and of questions about knowledge asked from the examiner situation makes this plausible. If we adopted the metaepistemological picture under discussion, we would face strong pressure to explain how those debates are relevant to our non-epistemological inquiries, or to abandon them as misguided.

Bibliography

- Aarnio, Maria Lasonen. Forthcoming. "Higher-Order Evidence and the Limits of Defeat." *Philosophy and Phenomenological Research* .
- Adler, Jonathan. 1981. "Skepticism and Universalizability." *The Journal of Philosophy* 78:143–156.
- Adler, Jonathan E. 1999. "The Ethics of Belief: Off the Wrong Track." *Midwest Studies in Philosophy* 23:267–285.
- . 2002. "Akratic Believing?" *Philosophical Studies* 110.
- Alexander, Joshua, Cameron Buckner, Chad Gonnerman, and Jonathan M. Weinberg. forthcoming. "Are Philosophers Expert Intuiters?" *Philosophical Psychology* 23:331–355.
- Alston, William, 2005. *Beyond "Justification": Dimensions of Epistemic Evaluation*. Cornell University Press.
- Alston, William P. 1980. "Level Confusions in Epistemology." *Midwest Studies in Philosophy* 5(1):135–150.
- . 1988. "The Deontological Conception of Epistemic Justification." *Philosophical Perspectives* 2:257–299.
- . 1993. "Epistemic Desiderata." *Philosophy and Phenomenological Research* 53:527–551.
- Antony, Louise. 2004. "A Naturalized Approach to the a Priori." *Philosophical Issues* 14:1–17.
- Armstrong, D. M., 1973. *Belief, Truth and Knowledge*,. London, Cambridge University Press.
- Arpaly, Nomy, 2003. *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford University Press.
- Austin, John Langshaw. 1956. "A Plea for Excuses." *Proceedings of the Aristotelian Society* 57:1–30.
- Berker, Selim. 2008. "Luminosity Regained." *Philosophers' Imprint* 8:1–22.

- Blackburn, Simon, 1984. *Spreading the Word*. Oxford: Clarendon Press.
- , 1993. *Essays in Quasi-Realism*. Oxford University Press.
- , 1998. *Ruling Passions*. Oxford: Oxford University Press.
- BonJour, Laurence, 1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- Brandom, Robert, 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, Ma.: Harvard University Press.
- , 2000. *Articulating Reasons: An Introduction to Inferentialism*. Harvard University Press.
- Brink, David, 1989. *Moral Realism and the Foundations of Ethics*. Cambridge University Press.
- Broome, John. 1999. "Normative Requirements." *Ratio* 12:398–419.
- Buckwalter, Wesley, and Stephen Stich. 2010. "Gender and Philosophical Intuition." *SSRN eLibrary* .
- Burnyeat, Miles, 1983. "Can the Skeptic Live His Skepticism?" In Burnyeat, Miles, editor, "The Skeptical Tradition," Berkeley, Ca.: University of California Press.
- Carnap, Rudolf, 1947/1956. *Meaning and Necessity*. Chicago: University of Chicago Press, second edition.
- . 1950. "Empiricism, Semantics, and Ontology." *Revue Internationale de Philosophie* 4:20–40. Reprinted in Carnap 1947/1956.
- Chalmers, David. Forthcoming. "Verbal Disputes." *The Philosophical Review* .
- Chalmers, David John, David Manley, and Ryan Wasserman, 2009. *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press.
- Christensen, David, 2007. "Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals." In Gendler, Tamar Szabó, and John Hawthorne, editors, "Oxford Studies in Epistemology, Volume I," pages 3–31. Oxford University Press.
- Cohen, Stewart. 1987. "Knowledge, Context, and Social Standards." *Synthese* 73.
- Conee, E., and R. Feldman. 1998. "The Generality Problem for Reliabilism." *Philosophical Studies* 89:1–29.
- Conee, Earl, and Richard Feldman, 2004. *Evidentialism*. Oxford: Oxford University Press.
- Craig, Edward, 1990. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Clarendon Press.

- Cuneo, Terence, 2007. *The Normative Web: An Argument for Moral Realism*. Oxford University Press.
- Darwall, S., 1983. *Impartial Reason*. Cornell University Press.
- Davidson, Donald, 1980. *Essays on Actions and Events*. Oxford University Press: Oxford.
- . 1982. “Rational Animals.” *Dialectica* 36:317–28.
- Dennett, Daniel, 1978. *Brainstorms*. MIT Press.
- , 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. Oxford University Press.
- , 1987a. *The Intentional Stance*. Cambridge, Ma.: The MIT Press.
- , 1987b. “Three Kinds of Intentional Psychology.” In “The Intentional Stance,” MIT Press.
- DePaul, Michael R., and Linda Trinkaus Zagzebski, 2003. *Intellectual Virtue: Perspectives From Ethics and Epistemology*. Oxford University Press.
- DeRose, Keith. 1995. “Solving the Skeptical Problem.” *The Philosophical Review* 104:1–52.
- Descartes, Rene, 2004. *Meditations on First Philosophy*. Caravan Books.
- Dretske, Fred. 2000. “Entitlement: Epistemic Rights without Epistemic Duties?” *Philosophy and Phenomenological Research* 60:591–606.
- Dretske, Fred I. 1970. “Epistemic Operators.” *Journal of Philosophy* 67:1007–1023.
- , 1981. *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press.
- Egan, Andy. 2008. “Seeing and Believing: Perception, Belief Formation and the Divided Mind.” *Philosophical Studies* 140:47 – 63.
- Elga, Adam. 2000. “Self-Locating Belief and the Sleeping Beauty Problem.” *Analysis* 60:143–147.
- . 2007. “Reflection and Disagreement.” *Nous* 41:478–502.
- Feldman, Richard. 1981. “Fallibilism and Knowing That One Knows.” *Philosophical Review* 90:266–282.
- , 2001. “Voluntary Belief and Epistemic Evaluation.” In Steup, Matthias, editor, “Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue,” Oxford University Press.

- . 2005. “Respecting the Evidence.” *Philosophical Perspectives* 19(1):95–119.
- . 2008. “Modest Deontology in Epistemology.” *Synthese* 161.
- Feldman, Richard, and Ted Warfield, editors, 2010. *Disagreement*. Oxford University Press.
- Field, Hartry, 2005. “Recent Debates about the A Priori.” In “Oxford Studies in Epistemology, Volume I,” Oxford University Press.
- Foley, Richard, 2001. *Intellectual Trust in Oneself and Others*. Cambridge University Press.
- Gasking, Douglas, 1962. “Avowals.” In “Analytical Philosophy I,” Basil Blackwell.
- Gendler, Tamar Szabó. 2008. “Alief and Belief.” *Journal of Philosophy* 105:634–663.
- Gibbard, Allan, 1990. *Wise Choices, Apt Feelings*. Cambridge Massachusetts: Harvard University Press.
- , 2003. *Thinking How to Live*. Cambridge, Ma.: Harvard University Press.
- Goldman, Alvin, 1979. “What is Justified Belief?” In Pappas, George S., editor, “Justification and Knowledge,” Reidel.
- Greco, Daniel. Forthcoming. “The Impossibility of Skepticism.” *The Philosophical Review* .
- Haidt, Jonathan, 2003. “The Moral Emotions.” In R. J. Davidson, K. R. Scherer, and H. H. Goldsmith, editors, “Handbook of Affective Sciences,” Oxford University Press.
- Hawthorne, John, 2004. *Knowledge and Lotteries*. Oxford University Press.
- Hawthorne, John, and Ofra Magidor. 2009. “Assertion, Context, and Epistemic Accessibility.” *Mind* 118:377–397.
- Heal, Jane. 1978. “Common Knowledge.” *Philosophical Quarterly* 28:116–131.
- Hemp, David, 2006. “The KK (Knowing that One Knows) Principle.” In “Internet Encyclopedia of Philosophy,” <http://www.iep.utm.edu/kk-princ/>.
- Hintikka, Jaakko, 1962. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press.
- Horwich, Paul. 1991. “On the Nature and Norms of Theoretical Commitment.” *Philosophy of Science* 58:1–14.
- Howson, Colin, and Peter Urbach, 1996. *Scientific Reasoning: The Bayesian Approach*. Open Court.

- Hume, David, 1740/2006. *An Enquiry Concerning Human Understanding*. Oxford: Clarendon Press.
- , 2000. *An Enquiry Concerning Human Understanding: A Critical Edition*. Oxford: Oxford University Press.
- , 2007. *Treatise of Human Nature*. Clarendon Press.
- Hurley, Susan, 1989. *Natural Reasons*. Oxford University Press.
- James, William, 1948. *Essays in Pragmatism*. New York, Hafner Pub. Co.
- Johnston, Mark. 1992. “How to Speak of the Colors.” *Philosophical Studies* 68:221–63.
- Kaplan, Mark. 1985. “It’s Not What You Know That Counts.” *Journal of Philosophy* 82:350–363.
- Kelly, Thomas, 2010. “Peer Disagreement and Higher Order Evidence.” In “Disagreement,” Oxford University Press.
- Kolodny, Niko. 2005. “Why Be Rational?” *Mind* 114:509–563.
- Kripke, Saul, Forthcoming. “On Two Paradoxes of Knowledge.” In “Collected Papers, Volume 1,” Oxford University Press.
- Lehrer, Keith, 2000. *Theory of Knowledge*. Westview Press.
- Lewis, David. 1979. “Scorekeeping in a Language Game.” *Journal of Philosophical Logic* 8:339–359.
- . 1982. “Logic for Equivocators.” *Noûs* 16:431–441.
- . 1986. *On the Plurality of Worlds*. London: Blackwell.
- . 1999. “Elusive Knowledge.” In “Papers in Metaphysics and Epistemology,” Cambridge University Press.
- . 2002. *Convention: A Philosophical Study*. Blackwell.
- Littlejohn, Clayton. Ms. “The Myth of the False, Justified Belief.” .
- Lycan, William G. 1985. “Epistemic Value.” *Synthese* 64.
- Macfarlane, John, Forthcoming. “Relativism and Knowledge Attributions.” In Bernecker, Sven, and Duncan Pritchard, editors, “Routledge Companion to Epistemology,” Routledge.
- Marušić, Berislav. Forthcoming. “The Self Knowledge Gambit.” *Synthese* .
- Melchert, Norman. 1985. “Why Constructive Empiricism Collapses Into Scientific Realism.” *Australasian Journal of Philosophy* 63:213–215.

- Millikan, Ruth G., 1984. *Language, Thought and Other Biological Categories*. MIT Press.
- Millikan, Ruth Garrett. 1989. "Biosemantics." *The Journal of Philosophy* 86:281–297.
- Myers-Schulz, Blake, and Eric Schwitzgebel. Forthcoming. "Knowing that P Without Believing that P." *Noûs* .
- Nagel, Jennifer. Forthcoming. "Intuitions and Experiments: A Defense of the Case Method in Epistemology." *Philosophy and Phenomenological Research* .
- Neta, Ram, and Guy Rohrbaugh. 2004. "Luminosity and the Safety of Knowledge." *Pacific Philosophical Quarterly* 85:396–406.
- Nozick, Robert, 1981. *Philosophical Explanations*. Harvard University Press.
- Owens, David. 2002. "Epistemic Akrasia." *The Monist* 85.
- Pettit, Philip, and Michael Smith. 1996. "Freedom in Belief and Desire." *Journal of Philosophy* 93:429–449.
- Pollock, John, 1974. *Knowledge and Justification*. Princeton, NJ: Princeton University Press.
- Prinz, Jesse, 2007. *The Emotional Construction of Morals*. Oxford University Press.
- Pryor, James. 2000. "The Skeptic and the Dogmatist." *Noûs* 34.
- Quine, Willard Van Orman. 1951/1953. "Main Trends in Recent Philosophy: Two Dogmas of Empiricism." *The Philosophical Review* 60:20–43. Reprinted, with alterations, in Quine 1953/1980.
- , 1953/1980. *From a Logical Point of View*. Cambridge, MA: Harvard University Press, revised second edition.
- Quine, W.V.O., and J.S. Ullian, 1978. *The Web of Belief*. McGraw-Hill.
- Raz, Joseph. 2007. "Reasons: Practical and Adaptive." *SSRN eLibrary* .
- Scanlon, Thomas, 1998. *What We Owe to Each Other*. Cambridge Massachusetts: Belknap Press.
- Schafer, Karl. Unpublished Manuscript. "The Varieties of Epistemic Evaluation." .
- Schwitzgebel, Eric. 2001. "In-Between Believing." *Philosophical Quarterly* 51:76–82.
- . 2002. "A Phenomenal, Dispositional Account of Belief." *Nous* 36:249–275.
- Shoemaker, Sydney. 1994. "Lecture II: The Broad Perceptual Model – Self Knowledge and "Inner Sense"." *Philosophy and Phenomenological Research* 54(2):271–290.

- Smithies, Declan. Forthcoming. "Moore's Paradox and the Accessibility of Justification." *Philosophy and Phenomenological Research* .
- Sorensen, Roy A., 1988. *Blindspots*. Oxford University Press.
- Sosa, Ernest. 2007. "Experimental Philosophy and Philosophical Intuition." *Philosophical Studies* 132:99–107.
- Stalnaker, Robert, 1984. *Inquiry*. Cambridge, MA: The MIT Press.
- . 1991. "The Problem of Logical Omniscience, I." *Synthese* 89.
- , 1999. *Context and Content*. Oxford: Oxford University Press.
- Stampe, Dennis W. 1977. "Towards a Causal Theory of Linguistic Representation." *Midwest Studies in Philosophy* 2:42–63.
- Stanley, Jason, 2005. *Knowledge and Practical Interests*. Oxford University Press.
- Stanovich, Keith E., and Richard F. West. 2000. "Individual Differences in Reasoning: Implications for the Rationality Debate?" *Behavioral and Brain Sciences* 23:645–665.
- Stine, Gail. 1971. "Dretske on Knowing the Logical Consequences." *Journal of Philosophy* 68:296–299.
- Stroud, Barry, 1984. *The Significance of Philosophical Skepticism*. Oxford: Oxford University Press.
- van Fraassen, Bas, 1980. *The Scientific Image*. Oxford: Oxford University Press.
- Vanderschraaf, Peter, and Giacomo Sillari, 2001. "Common Knowledge." In "Stanford Encyclopedia of Philosophy," <http://plato.stanford.edu/entries/common-knowledge/>.
- Vogel, Jonathan. 2000. "Reliabilism Leveled." *Journal of Philosophy* 97:602–623.
- . 2007. "Subjunctivitis." *Philosophical Studies* 134:73 – 88.
- Weatherson, Brian. 2004. "Luminous Margins." *Australasian Journal of Philosophy* 82:373 – 383.
- . 2008. "Deontology and Descartes's Demon." *Journal of Philosophy* 105.
- Wedgwood, Ralph. 2011. "Justified Inference." *Synthese* .
- Weinberg, Jonathan M., Shaun Nichols, and Stephen Stich. 2001. "Normativity and Epistemic Intuitions." *Philosophical Topics*, 29:429–460.
- White, Roger. 2005. "Epistemic Permissiveness." *Philosophical Perspectives* 19:445–459.

Williams, Bernard, 1970. "Deciding to Believe." In "Problems of the Self," Cambridge University Press.

Williamson, Timothy, 2000. *Knowledge and its Limits*. Oxford University Press.

———. 2001. "Comments on Michael Williams' Contextualism, Externalism and Epistemic Standards." *Philosophical Studies* 103:25 – 33.

———, 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.

———. 2011. "Philosophical Expertise and the Burden of Proof." *Metaphilosophy* 42:215–229.