Contextualising Knowledge: Epistemology and Semantics

Jonathan Jenkins Ichikawa

June 2, 2016

Contents

Acknowledgements

The kernels of this book were part of the Ph.D. dissertation I wrote at Rutgers in 2008 under the supervision of Ernest Sosa. I was thinking at the time about imagination and its many implications for epistemology—among those implications were the mental manipulations of possibilities, and the relationships between far-fetched possibilities, evidence, and skepticism. One strand of thought from my dissertation developed into a research program on thought experiments, intuitions, and the a priori. My first book, *The Rules of Thought* (OUP 2013), co-authored with Benjamin Jarvis, was a product of that strand. Another central strand in my dissertation had to do with knowledge, evidence, and relevant alternatives. The present book is a development of that project.

As is no doubt obvious, I have been enormously influenced by Timothy Williamson's *Knowledge and Its Limits* (OUP 2000) and David Lewis's 'Elusive Knowledge' (*Australasian Journal of Philosophy* 74(4), 549–67, 1996). Digesting those works has provided me enormous philosophical nourishment over the past several years; this book is my attempt to share the result.

I a grateful to many philosophers and friends who have had useful feedback and conversations with me about the material in this book. In many cases, the book has improved directly as a result. (Many others are prompts for possible future projects.) Some of these people are: Mark Alfano, Maria Alvarez, Nomy Arpaly, Steve Bahnaman, Derek Baker, Derek Ball, Aleksey Balotskiy, Alexander Bird, Liam Kofi Bright, Jessica Brown, Yuri Cath, Nathan Cockram, Stewart Cohen, Roger Clarke, Paul Dimmock, Travis Dumsday, Jeremy Fantl, Claire Field, Daniel Fogal, Rachel Fraser, Mikkel Gerken, Jasper Heaton, Torfinn Huvenes, Alex Jackson, Benjamin Jarvis, Masashi Kasaki, Carly Kocurek, Jennifer Lackey, Maria Lasonen-Aarnio, Karen Lewis, Lauren Leydon-Hardy, Sarah Little, Clayton Littlejohn, Dustin Locke, Errol Lord, Ned Markosian, Aidan McGlynn, Matt McGrath, Robin McKenna, Rachel McKinnon, Lisa Miracchi, Veli Mitova, Chris Mole, Adam Morton, Jennifer Nagel, Shyam Nair, David Plunkett, Kathryn Pogin, Lewis Powell, Baron Reed, Brian Renne, Katernya Samoilova, Jonathan Schaffer, Josh Schechter, Susanna Schellenberg, Eric Schwitzgebel, Daniel Star, Chris Stephens, Daniele Sgaravatti, Ernest Sosa, Jason Stanley, Kurt Sylvan, Evan Thompson, Shelley Tremain, Servaas van der Berg, Brian Weatherson, Nathan Weston, John Wigglesworth, Timothy Williamson, Chase Wrenn, Crispin Wright, and Sarah Wright. Thanks to L. Syd Johnson for a drawing of a lion. I am especially grateful to Nate Bemis, Carrie Ichikawa Jenkins, and two anonymous OUP referees, who each gave me thorough feedback on an earlier draft of the whole manuscript. There are many more people who ought to be on the list; I regret that I wasn't organised and thoughtful enough to remember to include them. (Thank you, and I'm sorry.) Thanks also to Peter Momtchiloff for all of his help with the transition from prose to print.

I presented work in progress on various parts of this book at many conferences, workshops, and colloquia. Thanks to helpful audiences at the 2009 Arché-Logos conference, the 2010 and 2012 Joint Sessions of the Aristotelian Society and Mind Association, the 2011 meeting of the Canadian Society for Epistemology, the 2012 WCPA meeting, the 2013 APA Central division meeting, the 2013 'Gettier Problem at 50' conference in Edinburgh, a 2015 epistemology conference in Kyoto, the 2015 'Factive Turn' conference in Vienna, and the 2016 CPA Annual Congress. Thanks also

to workshop participants at UBC, St Andrews, Helsinki, Northwestern, colloquium audiences at Indiana–Bloomington, British Columbia, Victoria, SUNY-Fredonia, CEU, and Lingnan, and seminar participants at Arché and the Edinburgh Epistemology Reading Group.

Preliminary work on this book was done at the Arché Philosophical Research Centre at the University of St Andrews, while I was part of an AHRC-funded project on intuitions and philosophical methodology; the bulk of it was written at the University of British Columbia, funded in part by a Hampton Grant.

The material in this book is published here for the first time, but some of the philosophical ideas derive from ones given in previously published papers (often with changes in emphasis, and occasionally with changes in view); this will be indicated along the way. Particularly relevant papers are:

- 'Quantifiers, Knowledge, and Counterfactuals', *Philosophy and Phenomenological Research*, 82 (2), March 2011: 287–313.
- 'Quantifiers and Epistemic Contextualism', *Philosophical Studies*, 155 (3), June 2011: 383–98.
- 'Knowledge Norms and Acting Well', *Thought*, 1, March 2012: 49–55.
- 'Basic Knowledge and Contextualist "E=K"', *Thought*, 2 (4), December 2013: 282–92.
- 'Justification is Potential Knowledge', *Canadian Journal of Philosophy*, 44 (2), July 2014: 184–206.

Introduction

I begin with an analogy.

Diary of a Narcissist

Here is a diary entry by Reginald, a confused narcissist.

> Dear Diary,
>
> I am perturbed. As you know, I've long found myself, if not perhaps perfection, surely the next best thing to it; I thank Providence every day for raising me up so far above the vulgar. It is no exaggeration to say that hitherto, I have counted myself among the very most beautiful and significant people in all of Creation. But today I happened across a paper by a philosopher called David Kaplan. What I found there shook my deepest convictions to the core. For Kaplan argued that certain words—'demonstratives' or 'indexicals', he calls them—are *context sensitive*; the referents of these terms can *vary* according to the conversational context in which they're used. My first instict was a sanguine one; his seemed an interesting and plausible semantic thesis. The referent of the word 'that', for example, is simply whatever it is at which my flawless finger happens to be pointing at a given moment when I'm speaking.
>
> But it's not just that! It's one thing to recognise the general semantic framework—it's quite another to make particular entries in the list of context-dependent terms. Among Kaplan's list of context-dependent terms are the very dearest and most important to me! He includes on his list, for example, such touchstones as 'I' and 'me'! Can you imagine, diary? *I*—Reginald the all-right—dependent on such contingencies as conversational contexts? Never in my wildest dreams would I have imagined that anyone would so trivialise *me*. Needless to say, I am deeply shaken. Must I accept that I am so unimportant? That there is nothing special about me, but rather than I'm merely, *whoever* happens to be speaking in a given conversation? The thought terrifies me. Tomorrow I shall attempt to rebut Kaplan's defamatory arguments; tonight I am too shocked. I must rest.
>
> Fondly,
>
> Reginald

Reginald's error is not difficult to diagnose; he's very bad at the use–mention distinction. At times, in the passage above, he is *using* words like 'I' and 'me', thus talking about himself; at other times he's *mentioning* them, thus talking about those English *words*. (You are a person, not an English word; 'you' is an English word, not a person.) Kaplan (1977) gives a context-sensitive theory about the *word* 'I'; but in doing so, he doesn't give a theory about me or Reginald or anybody else. Saying

that 'I' just picks out whoever happens to be speaking isn't tantamount to saying that anybody is unimportant. Kaplan's theory of demonstratives does not imply that Reginald isn't special. So much is, I take it, pretty obvious.

Nevertheless, when epistemologists start thinking about contextualism about 'knows'—roughly, the thesis that 'knows' is similar to words like 'I' and 'me' in that its referent depends on the conversational context—the corresponding point is not always treated as quite so obvious. For example, echoing Reginald, I have often encountered a perceived tension between contextualism about 'knows' and the idea that knowledge is important. I encounter this perception more often in conversation than in print, but Alvin Goldman does give a brief expression to a version of it here:

> A popular view in contemporary epistemology (with which I have much sympathy) is that knowledge has an important context-sensitive dimension. The exact standard for knowledge varies from context to context. Since it seems unlikely that natural kinds have contextually variable dimensions, this renders it dubious that any natural kind corresponds to one of our ordinary concepts of knowledge. Goldman (2007, p. 17)[1]

Assuming that the 'popular view' in question is contextualism, Goldman's fallacy is the same as Reginald's: it is a use–mention error. There is no straightforward connection between the semantic properties of the English word 'knows' and the metaphysical properties of knowledge. (Compare the fact that there is no straightforward connection between Kaplan's observations about indexicals and Reginald's beliefs about himself.)[2]

This book is about the relationship between contextualism about 'knows', on the one hand, and epistemological theorising about knowledge, on the other. It is a mistake to think that there is any very straightforward connection between them, but is there a subtler one? I shall suggest that there is. In particular, I will argue that there is a mutually supporting package of views, combining a particular brand of contextualism about 'knows' with a particular interpretation of the 'knowledge first' programme, according to which knowledge is a theoretically fundamental and important mental state. Assuming contextualism, the sentence, 'knowledge is a theoretically fundamental and important mental state' may be a context-sensitive one—this sentence could be used to express different propositions in different conversational contexts. That doesn't mean it isn't true and informative and theoretically enlightening, or that we can only mention, rather than use it. But it does mean that when we use it, we must use it carefully, and attend to potential ambiguities.

I have written a book about the relationship between contextualism about 'knows' (hereafter 'contextualism') and the knowledge first programme because I think both of these views have much to commend them. I have defended versions of both views in print—many of the ideas from these earlier papers are incorporated into this book. Along the way, I will say something about why I find these two views attractive, and I hope that many readers will come to look sympathetically on them, but it is not my primary purpose to argue for either contextualism or the knowledge first stance. Rather, I hope to show that these disparate views, though independently developed and considered, and widely thought to stand in a kind of tension, in fact fit rather well together. Contextualism can help the knowledge-first theorist respond to certain important objections to that stance; knowledge-first can help provide a theoretical motivation for the contextualist's claims. Any philosopher who wants to consider one of these views would do well to do so along with the other. Throughout

[1] In work in progress, Dani Rabinowitz also defends a version of Goldman's argued incompatibility.

[2] Compare also the remarks of Jenkins (2008, p. 70): 'Note that none of this context sensitivity [about 'explanation' and related terms] gives us any reason to be mind-dependent anti-realists about explanation. What depends on contexts (and hence on the intentions, interests and so on of the utterer and/or audience) is what is expressed by terms like 'good explanation' and 'best explanation', *not* what counts as a good or bad explanation *once these aspects of meaning are settled*. The thought is also related to Keith DeRose's 'intellectualist' motivation for contextualism—see DeRose (2009, ch. 6). However, I will suggest in Ch. 1 that this motivation is not mandatory for contextualists.

the book I'll canvass a series of studies of particular issues, exploring particular ways in which contextualism and the knowledge first stance may synergise. I include a chapter summary at the end of this introduction. First, however, I should set out the important backdrop that is the knowledge first program.

Knowledge First

It's not particularly controversial that knowledge is epistemologically interesting; but in what way is it interesting? A prevalent assumption in some of the history of philosophy had it that knowledge is a central *explanandum* in epistemology: a central task is to explain what knowledge is, how or whether knowledge can be attained, whether and why knowledge is particularly valuable, etc. Paradigmatic of this approach is the "theory of knowledge" literature spawned by Edmund Gettier's famous paper. The aim of this literature was to provide an "analysis" of knowledge in more fundamental terms—to explain knowledge, for example, in terms of belief, justification, truth, evidence, etc.

More recently, some philosophers have attempted to approach epistemological questions concerning knowledge from a different angle. Timothy Williamson is widely credited with this change in perspective; his 2000 book, *Knowledge and Its Limits*, advocated for an approach he calls "knowledge first".[3] The knowledge first stance reverses the traditional order of explanation: knowledge is treated as *explanans*, rather than as explanandum. The idea here is that knowledge is in some sense fundamental, and can be used to illuminate other states of epistemological interest. For example, perhaps we should understand evidence in terms of knowledge, instead of vice versa.

In our (2016), Carrie Jenkins and I distinguished between a number of distinct knowledge-first theses. In particular, we argued for a clear distinction between the *metaphysical* and *representational* claims that travel under that banner. Metaphysical knowledge first claims are claims about knowledge itself; questions about whether knowledge is a mental state, or whether it is a (relatively or absolutely) fundamental feature of reality are foregrounded. Representational claims, by contrast, have to do with how we *think* about or *talk* about knowledge; questions about whether the concept KNOWS has BELIEVES as a component, or whether knowledge ascriptions typically or invariably proceed by virtue of belief ascriptions, for instance, characterise the issues in discussing representational knowledge first theses.

Although Williamson defends views of both families, and although some authors have argued for strong connections between them, Jenkins and I argued that the views are prima facie independent. The version of the 'knowledge first' programme I am interested in exploring in this book belongs to the *metaphysical* family. I am more interested, for example, in the idea that knowledge is a mental state that has theoretically significant roles in explaining things like action, belief, and justification, than I am in the idea that KNOWS is a mental state concept, or the question whether knowledge ascriptions proceed via tacit belief ascriptions. Throughout the book, I will be interested in exploring ways in which knowledge connects to other areas. My sympathies lie with the broad methodology of Lewis (1984), according to which appearance in good theorising about the world is a mark of fundamentality.

Lewis himself focused on the *perfectly* natural; his view is that the perfectly natural properties are the ones that appear in the basic elements of a minimally adequate theory of the world. In my view, it is reasonably natural to extend that thought to the idea that *relatively* fundamental properties appear in less fundamental, but genuinely real, theories.[4] The idea of the knowledge first project, so

[3] Contemporary pre-cursors included Zagzebski (1994), Zagzebski (1996), and Craig (1990). There were certainly earlier precedents, e.g. Kneale (1949, p. 18): 'According to the view presented here, knowledge is sui generis and the two varieties of belief are to be defined by reference to it.' For an overview and discussion of historical knowledge-first ideas, see Marion (2000) and Mulligan (2014).

[4] The same, I think, goes for Ted Sider's more recent treatment of similar themes—see Sider (2011, p. 29).

interpreted, is that knowledge is *more metaphysically fundamental* than, for example, was supposed in the project of trying to explain knowledge in terms of justification and belief. This metaphysical claim should be understood in a way similar to the way that many of us think that green is metaphysically prior to grue, or that electrons are metaphysically prior to laptop computers. The approach is plausible to the degree to which metaphysical theorising in terms of knowledge is fruitful. This is why, for example, Timothy Williamson's suggestion that citing knowledge typically makes for a better explanation than does citing belief, in the explanation for someone's action[5]; that it appears in the best theory of action explanation is some reason to think knowledge is of some theoretical significance, hence metaphysical fundamentality.[6] So likewise for other theoretical roles posited for knowledge—perhaps one need invoke knowledge in order to explain evidence, or justification, or mental content, etc.

If a picture along these lines is right, then finding such roles for knowledge will constitute evidence that knowledge is a relatively fundamental state. To be sure, the Lewisian picture about fundamentality is controversial; adjudicating the serious metaphysical question about the relationship between theoretical roles and fundamentality is well beyond my present scope.[7] Even setting such relations aside, the idea that knowledge connects in deep theoretical ways with action, justification, belief, evidence, etc. is itself of significant interest.

If contextualism is correct, then the sentence, 'knowledge is a relatively fundamental state' may itself be a context-sensitive sentence. (I'll explain contextualism in much more detail in Ch. 1.) Contextualists need to have something to say about how it is intended. As will emerge throughout the book, I think that different proposed theoretical roles for knowledge ought to be treated differently in this respect. There is a kind of general perception among epistemologists that a knowledge first stance fits poorly with contextualism. I think this is a mistake.

Contextualism 'Evading' Epistemology?

One preliminary worry has to do with the general relevance of contextualism to epistemology. In effect, it starts with the observations I made above about the independence of claims about 'knows' ascriptions from claims about knowledge, and concludes from this that insofar as we care about the latter, we ought to ignore the former. As Ernest Sosa puts it, '[c]ontextualism replaces a given question [about knowledge] with a related but different question [about "knows" ascriptions]'. Sosa (2000, p. 1) And it is not at all clear that an answer to the latter question will bear on the former. To adapt one of Sosa's own pithy examples, Patience might say to herself:

> I am very confident that people often utter truths when they say "Somebody loves me." But does anybody at all love me?[8]

In exactly the same way, one may be convinced by the truth of contextualism, and rest easy that people often utter truths when making knowledge ascriptions[9]; still, one might wonder whether anybody at all knows anything. (If you're having a hard time seeing how this could be an open question, let's stipulate that I'm writing in at least a moderately skeptical context now; the thought is that people utter truths with such ascriptions only in more lax contexts.) It is important to keep clear the distinction between *object language* discussion, that *uses* terms to talk about the world

[5] Williamson (2000, pp. 62–3)

[6] See Ichikawa and Jenkins (2016, §§2.1, 4) for further discussion.

[7] For some discussion, see Jenkins (2013).

[8] Sosa (2000, p. 3)

[9] As I use the term, knowledge ascriptions are sentences using the word 'knows'.

more broadly, and the *meta-language*, that *mentions* terms in order to discuss them.[10] So it is that Sosa raises the challenge of how it is that contextualism can be relevant for epistemology. Hilary Kornblith has gone so far as to argue that contextualism is 'largely irrelevant to epistemological concerns'.[11]

Contextualism has two kinds of responses available to this worry. One of them is the thesis of this book: that a holistic theory combining contextualism and the uncontroversially epistemologically-relevant knowledge first programme is appealing; if contextualism helps make that programme more plausible, it is contributing helpfully to epistemology.

But there is also a more schematic reply available, which I'd like to articulate now. I agree with Sosa's claim that contextualism doesn't bear in any direct way on standard epistemological questions about the nature and extent of human knowledge. But I dispute the inference from this point to its *irrelevance* for the latter. For even though it doesn't provide any straightforward evidential support for any particular epistemological view, contextualism is, if true, crucially important for the *methodology* of epistemology. Anyone interested in understanding knowledge has an interest in thinking clearly about knowledge, and if contextualism is true, then equivocation on 'knows' is possible. So if contextualism is true, epistemologists must exhibit sensitivity to this fact.[12]

Take for example a classical skeptical argument like this one: I have no way to tell whether or not I will unexpectedly drop dead tonight; therefore I don't know much at all about what I may or may not do tomorrow. Its premise enjoys some intuitive plausibility, but its conclusion is far more skeptical than most epistemologists want to admit. It is tempting to suppose that one has to choose between these attractive ideas; then non-skeptics are burdened with the task of explaining away the attractiveness of the initial claim of ignorance. But if contextualism is right, then one needn't reject either intuitive starting-point; the argument to the effect that one must is equivocal. Careful attention to the language we use is sometimes the only way to avoid confusion about that which our language is about.[13]

This is in effect the same observation that Timothy Williamson makes in a different context in *The Philosophy of Philosophy*:

> Philosophers who refuse to bother about semantics, on the grounds that they want to study the non-linguistic world, not our talk about that world, resemble scientists who refuse to bother about the theory of their instruments, on the grounds that they want to study the world, not our observation of it. Such an attitude ... produces crude errors. Williamson (2007, pp. 284–85)

I do not wish to deny that there is also an important sense in which contextualism is not *the answer* to deep epistemological questions—I'll discuss this in much more depth in Ch. 3. But there is every reason to expect it to be *of relevance*.[14]

---

[10] The meta-language uses *some* terms—consider for example this sentence: "'Me' is an example of a context-sensitive term." This sentence *mentions*, rather than *uses* the word 'me', which is why it is about that word instead of me (i.e., Jonathan). But it also *uses* the word 'example' and six other words.

[11] Kornblith (2000, p. 24). As an OUP referee points out, in addition to the critique given in the main text, it seems that Kornblith's argument depends on the dubious (to understate things!) assumption that consideration of skepticism exhausts epistemology's concerns.

[12] See DeRose (2009, pp. 18–19) for a version of this point.

[13] Indeed, Sosa himself recognises as much in different contexts. Compare his remarks in Sosa (2009, p. 104) about a different kind of semantic issue: 'Semantic ascent does have a place in epistemology if only when we attempt to understand persistent disagreement by appeal to ambiguity or context-dependence.'

[14] Sosa also gives another argument against the epistemic significance of contextualism in his paper—one that relies on the assumption that contexts in which one engages in epistemology are inevitably skeptical ones relative to which all or nearly all knowledge ascriptions are false. I dispute this assumption in §6.2. See also Blome-Tillmann (2014, pp. 49–52).

Context-Sensitive Normative Discourse

A second preliminary objection to my project is more specific to the interaction between contextualism and knowledge norms. A significant component of the knowledge first stance relates knowledge to normative concepts. For example, Williamson (2000) argues that knowledge is the constitutive epistemic norm of assertion, which implies something like this:

(N) If and only if S knows *p*, S is epistemically permitted to assert that *p*

Other knowledge norms have also been proposed, many of which I will consider in detail in this book. But all of them, according to the objection I am now considering, fit badly with contextualism. The objection runs like this: any contextualist who adopts a knowledge norm is committed to problematic contextualism about normative concepts. Suppose, for example, that a contextualist endorses (N). Since the left-hand side is context-sensitive, by the contextualist's lights, our contextualist must also be a contextualist about the right-hand side too. But (the objection continues) it is not plausible to endorse this kind of normative contextualism in the cases at issue.[15]

The argument is fallacious; it semantically ascends and descends freely, and a contextualist need not accept it. Semantic ascent is the move from an object-language claim to a metalanguage claim, as in the move from the claim that somebody loves me to the claim that an utterance of 'somebody loves me' is true; semantic descent is a matter of *disquotation*, inferring an object-language claim from a metalanguage claim: the utterance of 'somebody loves me' is true, therefore somebody loves me. Neither move is generally valid for context-sensitive discourse. (Suppose that nobody loves me, but that some people love Patience. And suppose further that Patience is the person who says 'somebody loves me'. Then that sentence is true as uttered, even though nobody loves me.)

Returning to the case of (N), observe first that it is stated in the object language; it *uses* words like 'knows' and 'permitted'; it does not mention them. It is not *about* these words. (There are no quotation marks in (N).) Therefore the move from (N) and contextualism about the left-hand side to contextualism about the right-hand side is invalid. Only a strong metalinguistic generalisation of (N) could have these kinds of consequences.

In fact, there are at least five options for contextualists who accept object-language claims like (N):

1. Decline to endorse any metalinguistic principle, instead interpreting (N) as holding in a particular favored context.

2. Interpret the object-level claim as applying in S's context, as does DeRose (2009) about assertion.

3. Endorse the principle in full metalinguistic generality, positing context-sensitivity in the normative language, as per the objection suggested.

4. Endorse the principle in full metalinguistic generality, positing context-sensitivity in the relevant 'assert' language.

5. Endorse the principle in full metalinguistic generality, holding that the context-invariant right-hand-side applies any time the left-hand-side is true in any context.

[15] Williamson (2005b) and Hawthorne (2004, pp. 86–9) each give arguments in this neighbourhood against contextualism. While both these authors have subtle things to say about the relationship between contextualism and knowledge norms, I think both underestimate the extent of the contextualist's resources.

These represent a catalogue of available tools to the contextualist; there is no reason a contextualist need adopt a uniform treatment for all knowledge norms or other theoretical principles involving knowledge. I will apply diverse strategies through the book, applying a version of (2) to the knowledge norm of assertion, a version of (5) to certain proposed connections between knowledge and rational action, and a version of (4) to the equation of knowledge and evidence, and to the relationships between knowledge, justification, and belief.

The point of this section was to give one flavour of the kind of tension that some theorists have perceived between contextualism and knowledge-first epistemology. Several similar tensions (and some less similar ones) will be explored throughout the book.

Outline of the Book

Although contextualism and the knowledge first project have developed independently, and are typically thought to stand in tension with one another, I will argue that this perceived tension is illusory. On the contrary, I mean to make the case in this book for the idea that contextualism and the knowledge first project are complementary: each has something to offer the other. The details will come over the course of the book, but one common theme will be that contextualism helps the knowledge first project to avoid counterintuitive consequences, while the knowledge first project can help certain forms of contextualism answer the challenge of being ad hoc.

I begin in Chapter 1 with a development of the contextualist semantics I prefer. It is inspired by and related to David Lewis's relevant alternatives approach to knowledge ascriptions—it descends from an approach I first developed in Ichikawa (2011a). According to the Lewisian idea, the context-sensitivity of knowledge ascriptions is modelled on the context-sensitivity of modals or sentences involving quantifiers—satisfying 'knows *p*' requires evidence that conclusively rules out 'all' not-*p* cases, but where the domain of the 'all' depends in part on the conversational context. I'll argue in that chapter that a version of this idea captures quite a lot of the intuitive data, and avoids some of the challenges that have been levelled against contextualists. I'll also argue that there is a a significant class of under-explored linguistic data, concerning the interaction of knowledge ascriptions with conditionals, which further motivates this kind of contextualist approach.

Chapter 2 picks up on the idea that knowledge requires a certain kind of counterfactual connection to the truth. This rather natural idea played significant roles in twentieth-century theorising about knowledge, but is almost entirely discredited today—not least because it is thought to countenance 'abominable conjunctions'. However, the contextualist approach developed in Ch. 1, when combined with an independently-motivated contextualist approach to counterfactual conditionals, allows the proposed connection to be seen in a new light. I'll argue that sensitivity, suitably understood, *is* a genuine necessary condition for knowledge; given contextualism, counterintuitive consequences—including abominable conjunctions—can be avoided.

In Chapter 3, I turn to two pressing questions about evidence—one concerns Timothy Williamson's famous suggestion that a subject's evidence comprises all and only her knowledge ('E=K'). I agree with the letter of this equation (suitably embedded within contextualism); but I also suggest that there is good reason to recognise a privileged category of *basic* knowledge/evidence. This motivates the second central question of the chapter: what is the best way to understand basic evidence? The contextualist approach described in Ch. 1 had it that the truth of a knowledge ascription requires *evidence* that rules out all relevant counter-possibilities—but this could be understood in terms of various approaches to evidence itself. One of the under-recognised commitments of David Lewis's contextualism—and of quite a lot of contemporary philosophy—is a kind of 'Cartesian' approach to basic evidence, whereby a subject's subjective experiences are the things that are known first and best. I'll challenge this assumption in this chapter, arguing that it's certainly not mandatory, and

very plausibly ill-motivated. I'll consider instead a kind of neo-Moorean *disjunctivist* approach to evidence, according to which that which a subject can see for herself—including propositions about the external world—are basic evidence playing a foundational role. Consequently, the relationship between contextualism and radical skepticism ends up on my view rather different than that which most contextualists so far have posited.

Chapter 4 develops a knowledge-first theory of epistemic justification within the contextualist framework. Its starting point is the approach to justification defended in Ichikawa (2014), according to which justification is a matter of 'potential knowledge'—I characterised this notion as a matter of being intrinsically identical to a possible subject with knowledge. Ch. 4 generalises from that approach in two ways: first, it motivates and explicates a metasemantic generalisation of that view. The result is a contextualist semantics for ascriptions of epistemic justification to match that given for knowledge. Second, it relaxes the internalist assumption I'd previously made, considering various ways to understand the notion of 'potential knowledge', corresponding to various conceptions of basic evidence. One upshot of this generalisation is a new theoretical understanding of the kind of approach that could even motivate *radically* externalist theories of justification, such as the identification of justification with knowledge.

The final three chapters turn to proposed knowledge norms. Chapter 5 considers and defends the knowledge norm of practical reasoning, and suggests that, contrary to many authors' claims, it does not have radical externalist implications about rational action; relatedly, it also doesn't, when combined with contextualism, require any sort of implausible contextualism about normative discourse about action. I'll suggest that arguments to the contrary often tacitly assume a highly questionable approach to practical reasoning, which I call the 'thought bubble model'. By contrast, in Chapter 6, I *will* defend a kind of subject-relative version of the knowledge norm of assertion, justifying and systematising the connection by reference to Stalnakerian models for conversational contexts. Finally, Chapter 7 defends a contextualist semantics for belief ascriptions, and argues for a systematic normative relation between belief and knowledge, and between 'believes' and 'knows'.

Chapter 1

‘Knowledge’

What is contextualism about knowledge ascriptions, and why might someone want to be a contextualist?[1] This chapter lays out contextualism in general (§§1.1–1.4), and develops (§§1.5–1.7) the particular quantifier-based approach to ‘knows’ contextualism I favour. In §§1.8–1.11 I will go into some details on my preferred understanding of ‘epistemic standards’. (Readers interested only in the broad shape of the view and the connection to knowledge first epistemology might find it efficient to skip those sections.) I’ll conclude the chapter with a response to an objection to quantifier-based forms of contextualism from Jason Stanley (§1.12), a semantic motivation for the view (§1.13), and some remarks on the importance, within this framework, for a consideration of the theoretical roles of knowledge and their relations to contextualism (§1.14). The rest of the book will be devoted to the exploration of such connections.

1.1 Contextualism in General

Contextualism about a given bit of natural language is the view that that bit of natural language is context-sensitive. The minimal characterisation of contextualism I shall use is this: if a term is context-sensitive, then a sentence containing it will express distinct propositions in different conversational contexts.[2] It is easiest to introduce context-sensitivity by way of example. The most uncontroversially context-sensitive terms in English are the indexicals, for instance:

- I/you/he/she/him/her
- this/that/it
- here/there
- now/then

A given use of the word ‘she’ will refer to some person; but *which* person is referred to depends on the conversational context. If you ask me how Josephine is doing, and I utter, ‘she is sad’, I have

[1] As I use the term, *knowledge ascriptions* are sentences using the word ‘knows’. I’ll discuss my terminological conventions further in §1.2 below.

[2] A less minimal characterisation of the context-sensitivity of a term X would have it that X itself takes on different semantic values in different conversational contexts. I do not prefer this more committal account, in part because I think views like that of Schaffer (2004) about ‘knows’, or of Stanley and Szabó (2000) about quantifiers, are best counted as contextualist, even though they hold that ‘knows’ and quantifiers respectively always contribute the same semantic content. (They satisfy the more minimal statement of context-sensitivity because they posit tacit argument places which can be filled by conversational contexts in different ways.)

used those words to say that Josephine is sad. But if I'd uttered the very same words in a different conversational context—say, one in which you'd asked me how *Buttercup* was doing, those words would *not* have meant that Josephine is sad; they'd've meant that *Buttercup* is sad. So the sentence 'she is sad' might be about Josephine, or it might be about Buttercup, and it's the conversational context that determines which person it's about in a given instance. We can think of the person the sentence is about as the referent of a given instance of 'she'; since the referent depends on the context in which the word is spoken, contextualism about 'she' is true.

It is wholly uncontested that indexicals like 'she' are context-sensitive. It is only slightly more controversial to extend the treatment to gradable adjectives and quantifiers.[3] Gradable adjectives are terms that signify properties that come in degrees; the degree necessary to satisfy the semantic value of a given instance of the term will depend on the context in which the term appears. For example, 'rich' is a gradable adjective—there are many degrees of wealth, and some of them would suffice for 'rich' in the sailors' conversational contexts, but not in the Lord Commissioner's. Quantifiers are terms like 'all', 'some', and 'everybody'; these are typically understood to range over a domain of quantification that is determined in part by the conversational context. Consider, for instance, a sentence like 'everybody is singing'. Its truth requires that every member of a certain domain be singing; but which people are included in the domain—which people 'count' as part of 'everybody'—depends on the conversational context. In one conversation, it may include the twenty members of the opera company who are currently on stage; there, it is true, as all present are now singing. In another, it may include the baritone who is not on stage during this scene; there, its truth depends on whether he happens now to be singing in his dressing room. In a third context, the domain for 'everybody' may include the hundreds of people sitting quietly in the audience; there, it is false.

A Sophist might try to make this all sound very mysterious: how can it be that everybody is singing (here indicating the people on stage), *even though at the very same time not everybody is singing* (here indicating the audience)? If one is unfamiliar with context-sensitivity, and is not careful to distinguish the level of use from the level of mention, one might be bullied into supposing this to show that there are no absolute facts about whether everybody is singing, or that whether everybody is singing depends on which people one is thinking about, or that strictly speaking, one always speaks falsely when saying, 'everybody is singing'. But once we understand these basic tenets of context-sensitivity, the confusions implicit in this sort of reasoning become evident.[4]

1.2 Kaplan on Character and Content

David Kaplan has developed the orthodox framework for context-sensitive language. Kaplan (1977) distinguishes two aspects of meaning, which he calls *character* and *content*. The character of a word is its linguistic meaning—it is the sort of thing one might find in a dictionary; it's that of which one has tacit knowledge by virtue of being a competent speaker of a language. Returning to the example of the previous section, the character of the word 'she' is something like, 'the female who is salient in the context of the conversation'. The *content* of a word is its referent—in a sentence about Buttercup, the content of 'she' is Buttercup herself.[5] In the case of context-sensitive terms like 'she', the content varies according to context. Formally, character is understood as a function from

[3] What dissent there is comes from defenders of 'minimal semantics' like Cappelen and Lepore (2005) and Borg (2004). (Though note that even one of those authors since endorsed orthodoxy in this matter in Cappelen and Hawthorne (2009).)

[4] Unger (1975) argues that many gradable adjectives, like 'flat', actually have empty (or nearly-empty) extensions for just this reason. The standard diagnosis, with which I agree, is that Unger is mistaken in declining to invoke the context-sensitivity of these terms to resolve the apparent puzzles.

[5] Kaplan does not write univocally about content; he sometimes identifies contents with individuals, as I do in the main text, but sometimes describes them as intensions. See Braun (2015, §3.2) for discussion.

contexts to contents; contents are referents. For example, the character of 'I' is the function that returns the speaker in a given context; the content of a particular instance of the word is the person speaking.

With Kaplan, I will use 'context' to describe the relevant features of the speaker's conversational situation that play roles in determining contents for context-sensitive language. So the context will include things like the time of the utterance, the speaker, facts about what is being demonstrated, etc. If 'knows' contextualism is correct, then the context will also include something like an 'epistemic standard'—see §1.8 below. These are technical notions, being used in specific ways. These uses are reasonably well-established, albeit idiosyncratic relative to their ordinary uses. In particular, 'context' in colloquial English often just means 'situation'. So for instance a 'contextual theory of personality development' has it that the development of a subject's personality depends on the subject's environment.[6] This kind of approach is very different from a contextualist theory of knowledge ascriptions—it is not, for instance, any kind of semantic claim about language.[7] Throughout this book, I will be reserving 'context' for its distinctively linguistic use.

The usage I am adopting is, I think it's fair to say, standard in epistemological discussions of contextualism, but it hasn't always been so, and it isn't universally adopted even now.[8] One sometimes reads epistemologists discussing a view they call "subject contextualism", distinct from one they call "attributor contextualism". The latter is their name for what I am calling 'contextualism', but 'subject contextualism' uses 'context' in the colloquial way gestured at above: it emphasises features of the subject's situation. For example, Michael Williams's 'inferential contextualism' is a 'subject contextualist' view in that it holds that whether one's belief amounts to knowledge depends on certain features of one's inferential and dialectical situation—whether, for instance, one has had one's belief challenged in certain ways.[9] Similarly, David Annis's (1978) 'A Contextual Theory of Epistemic Justification' holds that whether a belief is justified depends on what kinds of doubts have been raised and by whom. I do not prefer to use the name 'contextualism' in a broad way that includes such approaches. Indeed, it is not clear that there is anything distinctive about the suggestion that the 'subject's context' is relevant for whether she has knowledge or justification. Rather, once we make explicit that by 'context' we mean only 'situation', that suggestion looks to be the mere truism that whether a subject has knowledge depends on the subject's situation. But of *course* that is true. For example, whether a subject knows p depends on whether her evidence in favour of p is misleading. No doubt, the so-called 'subject contextualists' intended to be expressing the more specific idea that *certain interesting features* of the subject's situation are relevant; but this is an extremely different kind of view than that of 'attributor contextualism'.[10] So I think it's with good reason that epistemologists tend not to use this language in this way any more. (The current standard term, 'subject-sensitive invariantism', at least does not suggest false similarities with contextualism, but in my view it's still not particularly clarifying; see §1.9.) I will therefore follow the convention deriving from the 1990s according to which 'contextualism' refers to a distinctively linguistic thesis.

Also following Kaplan, I will use 'index' to describe the situation relative to which a proposition is to be evaluated. For relatively simple engagement with relatively simple sentences like 'Ralph

[6] L'Abate and Bryson (1994, p. 3).

[7] Compare also Rachel McKinnon's invocation of the 'context-sensitivity' of assertability—as she makes clear, this is not related to the kind of 'context-sensitivity' at issue in contextualism. See McKinnon (2015, p. 58).

[8] Neta (2008a, p. 325) is one recent invocation of the 'subject contextualist' and 'attributor contextualist' language.

[9] See Williams (1991) and Williams (2001a). See Pritchard (2002) for helpful discussion and clarification of Williams's view (including the label 'inferential contextualism').

[10] Pritchard (2002, p. 115) argues that Williams's view deserves the name 'contextualist', not because it is subject contextualist—he observes as I do that there is nothing distinctive about that commitment—but because of a particular emphasis on particular features of the subject's situation. But I am not sure why an emphasis on those features represents a point of similarity with 'attributor-contextualism'. (In fact, the particular features he cites amount to *differences* between Williams's view and that of the 'attributor-contextualists' Pritchard discusses.)

is talking to her', the index, just like the context, will just be something like the actual world and time.[11] The context provides that 'she' refers to Josephine, so the sentence says that Ralph is talking to Josephine; the index is the actual world and time, and what is said is true just in case the actual world, at that time, is such that Ralph is talking to Josephine. But sometimes we will wish to consider whether what is said is true relative to *different* indexes. We might encode this into the language itself, by using an index-shifting operator—for example, if we say 'it is possible that Ralph is talking to her', the context provides Josephine for 'her' just as before, but now, we consider whether there is *any* possible world in which that content is true. Or we may consider the same question given the simpler sentence; when one says, in the actual world (talking about Josephine), 'Ralph is talking to her', one can consider, not only whether that really *is* true, but whether it *could* be true (is it true at *any* index?), or whether it *must* be true (is it true at *every* index?), or whether it *will* be true an hour from now (is it true at the index of the actual world and 10:13 p.m.?), or whether it *would* be true if Josephine had gone ashore this morning (is it true at nearest worlds in which she did?).

As highlighted in the Introduction, it is crucial, in thinking clearly about contextualism, to attend carefully to the distinction between use and mention. Far too often, discussions of contextualism are less than fully explicit on this point—indeed, the time it took to unravel 'subject contextualism' from contextualism may naturally be thought of as an instance of this difficulty.[12] Throughout this book, I shall endeavour to be as meticulous as possible. Unfortunately and inevitably, with this kind of rigour comes a degree of unloveliness. In my view, in an academic monograph of this sort, the aesthetic price is worth paying; I hope the reader will be convinced to agree. (At any rate, I beg the reader's indulgence.) I will follow the standard convention of using quotation marks to signify that a word is being mentioned; when I write 'knows' without quotation marks, I am *using* the term: I know that Keith DeRose is a contextualist who thinks 'knows' is context-sensitive. For further clarity, I will reserve 'say' for a relation to a proposition, and 'utter' for a relation to a sentence. I'll use 'knowledge ascriptions' to refer to the assertoric utterance of sentences using 'knows'.[13] So according to contextualism, what one says when one utters a knowledge ascription will depend on one's conversational context.

1.3 Modals

Before turning to 'knows', I'd like to consider one more example of context-sensitivity that does not have anything obviously closely to do with epistemology: modals.[14] Modals are terms like 'must',

[11] Whether time should be considered part of the index or not is contested. See e.g. Brogaard (2012) in favour of time-indexed propositions, and Partee (1973) against them. See also Schaffer (2012) for a more comprehensive overview of the debate. This controversy doesn't matter for my purposes.

[12] Other instances of this kind of shortcoming are manifest in different ways in Lewis (1996), discussed in §1.5 (see also Cohen (1998, fn 10, p. 292)), Buford (2009), discussed in 6.2, and DeRose (2002), discussed in §6.5. See also (McKinnon, 2015, p. 58) ('This is the heart of epistemic contextualism: whether an agent knows depends on the context.') and Reed (2010, p. 224) (Contextualism implies 'that whether someone can correctly be said to have knowledge can change abruptly, even when there is no change in her evidence or reliability.'). *Many* more examples could easily be added.

[13] It's convenient, if ungraceful, to apply these terms alike to positive sentences like 'Hebe knows that Joseph is hungry' and negative ones like 'Joseph doesn't know that Ralph plans to elope'; talk of 'knowledge ascriptions' is meant also to cover such cases of knowledge denials.

[14] There are relationships between modals and epistemology. One clear example concerns *epistemic* modals—the epistemic senses of 'might' (i.e., Josephine might be outside—I can't hear her sleeping) and knowledge. As many authors have observed, epistemic modals are naturally understood in terms of knowledge: what 'might' be the case is what is consistent with what is known. See e. g. von Fintel and Gillies (2007). Much of the debate about epistemic modals concerns *whose* knowledge is relevant here. A rather natural approach is a contextualist one like that of (Dowell, 2011); the dialectic between such views and competitors—especially relativist competitors like Egan et al. (2005)—is very rich, and quite similar to the parallel debate about 'knows'; but engaging it in any detail is beyond my present scope. For an overview and defence, see Dowell (2017). As Anderson (2014) points out, there is also flexibility with regard to *which* of a given subject's knowledge is relevant.

'may', 'necessarily', 'possibly', 'probably', 'ought', 'should', 'often', 'always', etc. In fact, it is not obvious that these *need* be construed as a separate category from those discussed above; one possible way—indeed, the way that I like—to understand modals is to think of them as quantifiers ranging over possible worlds. Like quantifiers generally, modals are context-sensitive. Consider a modal sentence like (1):

(1) Bill must inspect the winch.

We can understand the 'must' in this sentence as a universal quantifier over a contextually-influenced set of possible worlds. In some conversational contexts, (1) will receive a *deontic* reading, quantifying over those worlds in which Bill performs his duties—it says that in all such worlds, Bill inspects the winch. In other contexts, (1) might receive a nomic reading; here, the worlds in question might be worlds in which the laws of nature are as they actually are, and Bill's physical constitution is as it actually is. To say that 'Bill must inspect the winch' in such a context would be to say that these features entail that Bill inspects the winch. (Imagine a context in which the relationship between free will and determinism is being considered.) With a bit more imaginative work, we can even contemplate a conversational context in which the relevant set of worlds is the set of metaphysically possible worlds in the sense of Lewis (1986) or Kripke (1980); then (1) would say that in all metaphysically possible worlds, Bill inspects the winch. (This is not particularly plausible.) Or (1) can receive an epistemic reading, if its conversational participants are trying to work out what Bill's job on the ship must be; as their evidence rules out various hypotheses, they might eventually conclude that 'Bill must inspect the winch'—in all worlds compatible with their evidence, Bill inspects the winch.

We can think of these categories—deontic, nomic, metaphysical, epistemic—as *flavours* of modality. But even having settled on a flavour, there is a good case to be made that there is additional flexibility for context-sensitivity at play. Suppose, for instance, that Hebe has been told that Bill performs one of three jobs on the ship, and she's observed the other two jobs being carried out by other crewmen; it would not be at all surprising under such circumstances to hear her utter (1), and, supposing that the relevant appearances were not deceptive, there is a reasonably strong intuition that she could say something true by doing so. Nevertheless, there are possible conversational contexts in which she might proceed more cautiously. She's been *told* that Bill performs one of these three jobs, but isn't it *possible* that this was a lie? If such possibilities are considered, we might well say that (1) is false; that Hebe would say something true by uttering

(2) Bill might not inspect the winch.

Even the distinctively epistemic 'must', it seems, is subject to stricter and laxer standards. This brings us rather close to the 'knows'-contextualist's diagnosis of scepticism. Let us consider it now.

1.4 'Knows' Contextualism and Skepticism

To adopt contextualism about 'knows' is to hold that 'knows' is context-sensitive; that sentences involving it express different propositions in different conversational contexts. Consider sentence (3):

(3) Joseph knows that the Captain has a daughter.

According to contextualism, (3) will express distinct propositions in distinct conversational contexts. In one 'ordinary' context, (3) requires only that Joseph has some relevant familiarity with the

Captain's familial circumstances, but allows that he is a fallible human with epistemic limitations, while in another 'skeptical' context, (3) might require that Joseph be in a superhuman epistemic position—for example, it might require that Joseph be able to offer non-question-begging evidence against the hypothesis that he is the victim of a large conspiracy concerning the Captain's family, or even that he is a brain in a vat. If, as it has seemed plausible to many, it is impossible to attain such a standard, (3) will express an impossible proposition in such a skeptical context. A skeptic argues:

- Joseph has no evidence against this possibility: Joseph is a brain in a vat, and there is no Captain or Captain's daughter.
- If there is a possibility w such that proposition p is false, and subject S has no evidence against w, then S does not know that p. Therefore,
- Joseph does not know that the Captain has a daughter.[15]

The skeptic's argument is deeply controversial; both of its premises are contested.[16] I have little at present to add to the vast literature about whether the premises above can plausibly be rejected; I make only two observations. First, the skeptic's premises do enjoy a *prima facie* plausibility; epistemologists argue against them because they appear to lead to unacceptable conclusions, not because they are implausible on their faces. Second, the contextualist can accept both premises of this argument.[17] Given these two facts, contextualism does enjoy at least some intuitive advantage over non-contextualist views. The contextualist is able to accept the sceptic's claims, but to quarantine them—the argument stated above is sound, but its conclusion is given in context-sensitive language; its superficial appearance of overturning tenets of common sense (like the fact that Joseph knows that the Captain has a daughter) is illusory. This has been an important historical motivation for contextualism.[18]

For my own part, I find its treatment of skepticism compelling, and a good reason to accept contextualism. I shall spell out in more detail how I think skepticism and contextualism play out below. However contextualism *per se* is not *committed* to the truth of the skeptic's claims in her skeptical context. It *allows* them, but does not *require* them. If, for instance, we are convinced, on grounds independent of avoiding skepticism, that the first premise in the skeptic's argument is false—perhaps our best theories of perception and of evidence entail that perceptual content about the external world always counts as evidence—we could still accept contextualism. We would not, presumably, do so in order to avoid external-world skepticism, but this is not the only possible motivation for skepticism. DeRose (2009) argues at length that contextualism is the best way to explain ordinary linguistic use of the term 'knows'; the present work argues that contextualism forms an appealing unity with the knowledge first program. More moderate skeptical scenarios that allow direct knowledge of the external world via perception are also sufficient to generate similar puzzles. The conspiracy possibility mentioned above, and the lottery paradox, which I will discuss

[15] This is one of a variety of ways to spell out a skeptical paradox. Another would proceed more directly on the basis of an intuition to the effect that Joseph cannot know the skeptical scenario not to obtain and a kind of closure principle. See e.g. DeRose (1995). I don't think these difference in formulation matter for my immediate purposes.

[16] Mooreans like Sosa (1999) and Williamson (2000) deny the first premise; sensitivity theorists like Nozick (1981) and Dretske (1981) typically deny the second, thereby denying single-premise closure for knowledge. But I will argue in Chapter 2 that sensitivity theorists need not deny closure or the second premise.

[17] Note that so-called 'subject-sensitive invariantists' cannot accept the truth of both premises without succumbing to contextualism; neither premise includes any stipulation about whether or not a subject attends to the skeptical possibility, or has any pressing practical interest in whether the Captain has a daughter. (We may suppose that we are running the argument in a way isolated from Joseph's own attention.)

[18] See e.g. DeRose (1995), Lewis (1996).

below, are examples. (I'll discuss these issues, and whether we should think of moderate and radical skeptical scenarios as importantly different, in much more detail in Ch. 3.)

For now, I will take contextualism's treatment of skepticism as *prima facie* motivation for contextualism, and develop my preferred version of the view. My approach is heavily informed by David Lewis's; as we shall see, it is in many respects a development of his.[19] In §1.7, I'll go on to lay out the Lewisian view to set the stage for my development of it. In §1.8, I'll contrast the approach with that of the most prolific contemporary defender of contextualism, Keith DeRose. In the final sections of the chapter, I'll turn to a novel motivation for a contextualism of this kind, and consider some influential objections to the approach. This will set the framework for the synthesis of contextualism and a knowledge-first approach that I will pursue over the rest of the book.

1.5 Elusive Knowledge

According to David Lewis, knowledge ascriptions are best understood on the model of universal generalisations. Lewis gives us this gloss on 'S knows that *p*':

> S knows proposition P iff P holds in every possibility left uneliminated by S's evidence.
> Lewis (1996, p. 551)

This is not a statement of anything in the neighbourhood of contextualism; it is a claim about what it takes for a given subject to know a given proposition. The quoted sentence *uses* the English word 'knows', but it does not treat it as a subject matter. Consequently, it cannot be a statement of, or a commitment to, contextualism—a thesis about a certain English word, not about what it takes for a subject to know a proposition. It can be rather easy to fail to appreciate this point. (Indeed, Lewis himself ignores it throughout most of his paper; see DeRose (2009, pp. 212–17) for this complaint against Lewis.) But, as emphasised in the Introduction, contextualism implies that general object-level claims about knowledge are susceptible to various interpretations; Lewis intends this one to carry a certain kind of metasemantic generality—but one might sign up to this statement without embracing contextualism. Lewis himself wishes to exploit the context-sensitivity of the 'every possibility' quantifier in the right-hand side above. Which possibilities are relevant for the purpose of claims about 'every' possibility depends on the context of the speaker, and so do which possibilities one must consider to evaluate whether S satisfies 'knows *p*' in a context. Let's understand Lewis as signing up to some kind of claim like this one:

> 'S knows that *p*' is true in context C iff '*p* holds in every possibility left uneliminated by S's evidence' is true in C.[20]

As mentioned above, it is largely uncontroversial that quantifiers like 'every' are context-sensitive; Lewis attempts to characterise 'knowledge' discourse in a parallel way. In my first paper on contextualism, I characterised Lewis's strategy thus: 'Lewis's "knowledge" inherits its context-sensitivity from the context-sensitive "every possibility".'[21] I now consider this choice of words regrettable.

[19] The view of the next several sections was first laid out in Ichikawa (2011a). The core of that paper remains my current view, although the presentation of this book reflects some small changes, some additional commitments, and a difference in emphasis.

[20] Lewis (1996, p. 567) remarks that the conscientious reader who wants to understand his suggestion in a more precise way can perform the required 'semantic ascent' his or herself; this is the exercise I've engaged with in the previous paragraph. In some cases, however, this will be a more difficult task than Lewis seems to have anticipated. It is not at all clear, for instance, that there's *any* way, once we are using the language carefully and precisely, to make sense of Lewis's discussion of dogmatism on his pp. 564–5.

[21] Ichikawa (2011a, p. 385).

The word 'knowledge' does not contain or otherwise involve the phrase 'every possibility'; the context-sensitivity of the latter does not *generate* or *explain* context-sensitivity in the former. I do not think—nor, so far as I can tell, would Lewis have thought—that as a psychological description of humans, we tend to evaluate 'knowledge' statements by running through context-sensitive claims involving 'every'. (Note also that there is no privileged direction in the metasemantic claim above; the 'every possibility' claim needn't be getting at any kind of *essence* of the 'knows' claim.) Instead, I take Lewis's suggestion to be that we attribute context-sensitivity in 'knows' language *along the model of* context-sensitivity about quantifiers; the mere plausible truth of the metasemantic claim shows us a way to understand how 'knows' might behave in diverse conversational contexts. The familiar context-sensitivity of quantifiers provides us with a model for understanding the context-sensitivity of 'knows'.

Before pursuing the analogy further, let's first consider in some small detail contextualism about quantifiers.

1.6 Quantifiers

As indicated in §1.1, there are strong reasons to accept contextualism about quantifiers. Even ignoring the context-sensitivity of the gradable adjective 'cheap', what is said by an utterance of (4) depends on context:

(4) All the plates are cheap.

The Captain has been entertaining company in his cabin. In this context, an utterance of (4) makes a claim about the serving plates at his party. In fact, the insult is unfair; some of the Captain's plates are cheap, but some are rather dear.

As it happens, at the very same time, the boatswain is also hosting a party with some fellow sailors. The boatswain wants his friends to relax; in particular, he doesn't want them to worry about whether they might break any of the serving dishes. To this aim, the boatswain utters the same sentence: 'all the plates are cheap.' His use of (4) is true, even though the Captain's rude guest's was false.

Naturally, we're not too puzzled about (4) being true in one context and false in another, even though both utterances occur at just the same time, and just the same world; sentences like (4) are straightforwardly context-sensitive. *What is said* by an utterance of (4) depends on the context in which it is uttered. We can think of the character of the expression 'all the plates' as a (non-constant) function from contexts of utterance to properties of plates.[22] When the Captain's guest utters (4), the context supplies a property that picks out which plates are relevant—the property of being used at the Captain's party, for instance. When the boatswain utters (4), the context supplies a different property—the property of being used at the Boatswain's party, perhaps.

I say that the context interacts with the quantifier to produce a *property*. Why not say that the context determines a *set* of plates as the ones relevant for an utterance of 'all the plates'? Here are two reasons. First, consider the Captain's context. Suppose that there are four plates that are

[22] King (2005, p. 128) suggests that such complex expressions do not have Kaplanian characters, but rather that their contents are determined by contents of their simpler constituent parts. (Moreover, according to Stanley and Szabó (2000), an expression like 'all the plates' has among its contsituent parts a silent variable corresponding to a property.) Whether it is strictly correct to think of complex expressions as having substantive characters themselves doesn't really matter for my purposes; we can still describe a kind of lightweight character in the sense that the stable linguistic meaning of expression 'all the plates' identifies a function from conversational contexts to properties of plates. We may agree with King and Stanley, if we like, that this 'character' is explanatorilly idle, and that the real semantic work is done by the substantive characters of the constituent parts.

relevant, p_1–p_4. If the context provides a set of plates, then what is said by (4) in the Captain's context is a proposition about p_1–p_4. Now consider a counterfactual version of the Captain's context, in which a different four plates, say, p_5–p_8, are served at the party. Supposing that context fixes the set of relevant plates, in this counterfactual case, an utterance of (4) would have expressed a distinct proposition—one about p_5–p_8. So if the admiral utters (4), what is said would have been different, had p_5–p_8 been the plates present. But it is counterintuitive to claim that what is said depends on which plates happen to be around. Now this is admittedly less than a fully probative argument; there is nothing *incoherent* in the idea that what is said depends on one's environment, even in ways that the speaker would not recognise. Indeed, it's clear that some analogous cases hold: what is said by an utterance containing 'he', for example, depends on whether Tom or Dick is the person who happens to be present, whether or not the speaker is sensitive to the difference. Still, I think that there is a sense in which it sounds false to say that what is said (not merely whether it is true) in an utterance of (4) depends on facts like which plates the Captain has brought out; treating the context as picking out a property of plates, rather than a set of plates, lets us respect that intuition.

But even if you do not share my intuition here, there is also a stronger argument to the same conclusion. Consider the boatswain's context, where (4) expresses a truth. What he said by uttering (4) would have been false if the Captain had come to the party and brought some expensive plates. It's not only that an *utterance* of (4) would have been a *saying* of something false—it is that what was *actually said* would have been false. (Remember, as I'm using the terms, sentences are uttered; propositions are said.) The boatswain is saying a proposition that *is* true, but *would* have been false if there had been some expensive plates present. But then what is said must not have been a claim only about the plates that were actually used in the Boatswain's party. Suppose those were p_9–p_{11}, all of which are cheap. If context fixes a set of plates, then an utterance of (4) in the Boatswain's context picks out a proposition that is true just in case p_9–p_{11} are all cheap. But this proposition is true, not false, relative to an index in which the Captain crashes the boatswain's party and brings expensive plate p_3.[23]

This question—whether we should think of the contextual domain restriction of quantifiers as providing sets or properties—has a correlate in the case of the Lewisian approach to knowledge ascription: does the context provide a set of possibilities, or something more like a property of possibilities? If what I have said in this section about quantifiers is right, then the analogy suggests that it should be the latter. As I will lay out in §1.10, there are sound epistemological reasons to prefer this version of the view as well—I will suggest that the context provides a parameter that is best modelled as a function from indexes to sets of possibilities. First, however, let's return to the Lewisian approach to knowledge.

1.7 Lewis and Lewisian Contextualism

As indicated in §1.5, Lewis treats satisfaction of 'knows *p*' in a context to be equivalent in truth-value to a quantified claim about possible worlds in that context: satisfying 'knows *p*' in context C is a matter of eliminating 'all' not-*p* possibilities, where C also contributes to which possibilities 'count' as among 'all' the possibilities. Lewis goes on to offer a series of rules that are intended to determine, of a given possibility, whether it is relevant for the purpose of a given knowledge ascription. I shall rehearse Lewis's rules shortly; first, however, I should like to consider the role of these rules in Lewis's broader framework. I argued in Ichikawa (2011a) that the attempt to articulate the rules that govern whether a given possibility is relevant is a supererogatory one, for the purpose of the general Lewisian contextualist approach.

[23] See Stanley and Szabó (2000, p. 252) for a similar argument.

As mentioned above, contextualism about quantifiers is uncontroversial.[24] Which plates count as among 'all the plates' can vary according to the conversational context (as well as the index). We can support and clarify this claim by pointing to possible uses of 'all the plates' language that seem to require, for their intuitive truth conditions, that different plates are relevant in different contexts. It is *not* a commitment of contextualism about quantifiers—a view which enjoys a widespread orthodoxy—to articulate a series of rules that would determine, of a given plate, whether it counts as relevant, given a particular context and index. Natural language is an unwieldy beast; to demand precise rules that explain and predict the domains of its context-sensitive quantifiers is to demand too much. No doubt it would be nice to have them, but the plausibility of contextualism about quantifiers does not depend upon having identified them. It may be that the best that we can reasonably hope to do is to gesture somewhat vaguely at good rules of thumb: plates that are explicitly attended to in a context are relevant; plates that are in relevant respects similar to those explicitly attended to are relevant, etc.

In the same way, we should separate Lewis's ambitious attempt to characterise rules that determine which possibilities are relevant, given a particular context and index, from the more general statement of Lewisian contextualism sketched above, according to which satisfaction of 'knows p' requires ruling out those not-p possibilities that are relevant, given a particular context and index. The more general view can be motivated by uses of 'knows p' that require, for their intuitive truth values, that different possibilities be relevant in different contexts; it is further clarified by gestures towards what kinds of considerations contribute toward a given possibility's relevance or irrelevance. But the plausibility of Lewis's broader framework does not rest upon having articulated a set of rules that make all the right predictions.[25]

I'll use 'Lewisian contextualism' as a name for this thesis:

> *Lewisian contextualism.* The proposition expressed by an utterance of the form 'S knows that p' depends on the context in which it is uttered. In particular, the truth conditions require the elimination by S's evidence of all of some set of alternatives to p; which alternatives are relevant depends in part on the conversational context.

As I define it, Lewisian contextualism does not require the completion of the ambitious metasemantic project of articulating necessary and sufficient conditions for relevance. But Lewis himself does attempt this ambitious project, so it is worth saying something about his rules. He offers these four rules:[26]

[24] 'Uncontroversial' is a context-sensitive gradable adjective; I am writing in a context lax enough for this sentence to be true, despite the existence of *some* controversy, i.e. the position of Cappelen and Lepore (2005).

[25] Blome-Tillmann (2009) rejects some of the details of Lewis's rules, but retains Lewis's ambitious project of setting out rules that predict which possibilities are relevant; Ichikawa (2015) argues that Blome-Tillmann's attempt does not succeed, and that our ambitions should be lowered. One might interpret Blome-Tillmann's later emphasis on the one-direction 'simple view' (e.g. in Blome-Tillmann (2014) and Blome-Tillmann (2015a)), according to which 'S satisfies "knows p" in context C only if S's evidence eliminates all the $\sim p$-worlds compatible with what is presupposed in C,' as a move in the direction I suggest. (Still, I don't think it's the best one; notice also that, contrary to his remarks in Blome-Tillmann (2015a, pp. 2–3), the simple view does not itself ensure an 'interesting and important philosophical claim about the role of pragmatic presuppositions in the semantics of "knowledge"-attributions' or provide an 'explanation or resolution of skeptical puzzles' since, as a mere necessary condition for the satisfaction of 'knows', it allows that the role of presuppositions is idle. For example, the simple view is consistent with invariantist radical skepticism.) Even though they're not counterexamples to the simple view, the counterexaples given to the stronger version of Blome-Tillmann's view in Ichikawa (2015) would, if successful, demonstrate particular respects in which the simple view fails to explain what needs explaining.

[26] (Lewis, 1996, pp. 554–9). In fact, Lewis offers seven rules: the four mentioned in the main text, and three permissive rules that offer a defeasible presupposition that certain possibilities are irrelevant. (Lewis, 1996, pp. 558–9). As far as I can see, the shape of Lewis's view is the same whether or not the permissive rules are included—we can treat the restrictive rules as providing individually sufficient *and jointly necessary* conditions for relevance. So I will ignore the permissive rules. (To be frank, it's never been clear to me why Lewis included them in his list in the first place.)

- *Actuality.* The actual world is always relevant. Lewis clarifies that it is the *subject's* actuality—i.e., the world the subject is in—rather than the attributor's, that is at issue here.
- *Belief.* Any possibility the subject believes to obtain, or ought so to believe, is relevant.
- *Resemblance.* Any possibility that *saliently resembles* another relevant possibility (made relevant in some other way) is relevant.[27]
- *Attention.* Any possibility *attended to* in the attributor's context is relevant.

The rules of actuality and belief explicitly depend on the subject, rather than the attributor. In the Kaplanian terms, they are a matter of index, rather than context. The rule of attention, by contrast, has everything to do with the attributor, rather than the subject, so it is a matter of context instead of index. The rule of resemblance has a more complex, hybrid structure; the language of 'salience' suggests that which resemblance relation is at issue is a matter of context; but once that relation is fixed, which possibilities do resemble which others will be a matter of index.[28] I will take care in the following section to separate the respects in which the context is relevant from those in which the index is relevant; there is a sense in which it is misleading to list all four of these rules together. Or at least, it obscures a crucial feature of Lewisian contextualism.

Before moving on, I pause to note four more respects in which David Lewis's commitments in 'Elusive Knowledge' outstrip those of Lewisian contextualism, as I've characterised it. (Each is discussed in more depth in Ichikawa (2011a).)

First, Lewis signs up to a particular conception of possibilities: a possibility, for Lewis, is a centred (metaphysically) possible world. Consequently on Lewis's view there are no possibilities, and so no 'relevant' possibilities, in which, e.g. Hesperus is not Phosphorus. So it is impossible, by Lewis's own lights, to rule out the possibility that Hesperus is a star without ruling out the possibility that Phosphorus is a star. So anyone who knows the former must also know the latter; this is a counterintuitive result. But it is open to the proponent of Lewisian contextualism to construe possibilities more liberally. If, for example, there is a coherent notion of 'rational possibility' or 'conceptual possibility' that outstrips metaphysical possibility, there is nothing preventing the Lewisian contextualist from using this more liberal notion in his account. I think these *are* important notions—this is one of the central ideas in Ichikawa and Jarvis (2013)—and I prefer to invoke them in my own implementation. So I say that one's evidence might eliminate the possibility that Hesperus is a star without eliminating the possibility that Phosphorus is a star; in such cases, one might well know that Hesperus is not a star, but remain ignorant that Phosphorus is not a star. I do not plan essentially to be relying on such cases throughout the book; I merely record my own opinion on this controversial matter. If one prefers, with Lewis, to bite the bullet on cases like this, one need not quarrel with the project of this chapter.

Second, Lewis commits to a particular understanding of what it is for evidence to rule out a possibility. Evidence, Lewis says, comprises sensory and introspective experience; for some evidence to *rule out* a possibility is for the obtaining of the evidence—i.e., the fact that the subject has such-and-such experience—to entail that the possibility in question is non-actual. As Lewis puts it:

> When perceptual experience E (or memory) eliminates a possibility W, that is not because the propositional content of the experience conflicts with W. (Not even if it is the narrow content.) The propositional content of our experience could, after all, be false. Rather, it is the existence of the experience that conflicts with W: W is a possibility in

[27] The 'other way' proviso is to prevent the full recursive version of the view, which would threaten to extend to all possibilities via small steps of salient relevance.

[28] For more on the complexity of this feature of Lewis's view, see Cohen (1998) and Ichikawa (2011a, §8).

> which the subject is not having experience *E*. Else we would need to tell some fishy story of how the experience has some sort of infallible, ineffable, purely phenomenal propositional content ... Who needs that? Let *E* have propositional content *P*. Suppose even—something I take to be an open question—that *E* is, in some sense, fully characterised by *P*. Then I say that *E* eliminates *W* iff *W* is a possibility in which the subject's experience or memory has content different from *P*. I do *not* say that *E* eliminates *W* iff *W* is a possibility in which *P* is false. Lewis (1996, p. 553)

It is open to a defender of Lewisian contextualism to follow Lewis here, or to invoke a different understanding of what it would take for the evidence to rule out a hypothesis. For example, it is open to a Lewisian contextualist to invoke a richer understanding of 'evidence' than Lewis's own internalist conception; I will have much more to say about this topic (including how the approach fitswith a knowledge-first theory of evidence) in Chapter 3; for now, I'll leave open questions about the nature and extent of evidence.

Third, as several authors have emphasised, Lewis's Rule of Attention is too weak—it's not enough, for a possibility to be relevant, that it receive *any attention whatsoever*. One can hear a possibility mentioned without its becoming relevant. A more moderate version of Lewis's view would have it that if attributors *take a possibility seriously*, then it is relevant.[29]

Fourth, the Lewisian contextualist might embrace Lewis's general framework, according to which knowledge ascriptions require that alternatives among some contextually-influenced class of possibilities be eliminated by evidence, without following Lewis in supposing that this evidential state is *sufficient* for the truth of a knowledge ascription. Lewis embraced the surprising implication of his view that satisfaction of 'knows *p*' in a context does not require that a subject *believe* that *p*—'I even allow knowledge without belief, as in the case of the timid student who knows the answer but has no confidence that he has it right, and so does not believe what he knows' Lewis (1996, p. 556)—but this possibility is no commitment of the more general statement of Lewisian contextualism. In Ichikawa (2011a) I suggested adding a belief requirement and a basing requirement to the Lewisian framework. Here is my current proposal:

> S knows that *p* just in case, for some evidence E, (i) S believes that *p*, where that belief is properly based on evidence E, and (ii) all the E cases are *p* cases.

Strengthening the Lewisian condition for the truth of a knowledge ascription also avoids another problem for Lewis's own view: the implausible implication that the possession of evidence E is luminous, in Timothy Williamson's (2000) sense.[30] For suppose that some subject has evidence E. Then there are no possibilities (in the relevant set or otherwise) consistent with the subject's having that evidence, in which the subject does not have E. So regardless of what possibilities are included as relevant, 'knows that she has E' is satisfied. This is an implausible result; adding a belief and proper basing requirement avoids it.[31] So this is the approach I will continue to defend.

---

[29] See (Oakley, 2001) and Blome-Tillmann (2009, p. 247). Note that taking possibility *p* seriously doesn't require taking *every* possibility in which *p* seriously. Neta (2003, pp. 13–4) seems to assume the contrary, arguing that if one is considering the possibility that one doesn't have hands, one can't 'know' it not to obtain, since one can't rule out the possibility that one is a brain in a vat. But in the Lewisian framework, one may take the possibility that one doesn't have hands seriously by taking seriously the possibility that one was born without hands, or that one lost one's hands earlier in life. These possibilities are plausibly ruled out by one's perceptual experience; things would appear differently if they obtained.

[30] In fact, it has an even stronger implication than this already implausible one. If having evidence E is luminous, then any subject who has E is in a position to know that she has E. As I go on to explain in the main text, on Lewis's view, any subject who has E isn't merely *in a position* to know that she has E—she must *actually know* that she has E.

[31] Ichikawa (2011a) gave the characterisation above, minus the word 'properly'—it required only of the relevant E that the subject base a belief that *p* on E. I am now convinced that this was a mistake. It's not merely possible for one to fail to

1.8 Epistemic Standards

It is common to characterise contextualism in terms of 'standards' for knowledge or knowledge ascriptions. The notion is intuitive enough, but the extant literature about epistemic standards is unclear in crucial ways. Much of the confusion, I think, comes from the fact that the term is asked to cover two subtly, but importantly, different phenomena. As getting the correct notion of 'standards' right will have significant implications throughout the book, I will spend §§1.8–1.11 developing it. Readers interested in a quicker and broader treatment could skip ahead to §1.12 now.

To see why one might be tempted to explain contextualism about 'knows' in terms of standards, consider the oft-cited analogy of contextualism about 'knows' with contextualism about gradable adjectives like 'flat' or 'tall'.[32] The conversational context, it is natural to think, determines just how tall one has to be to satisfy 'tall' in that context. This threshold is very naturally thought of as a 'standard' for tallness. Just as a university might set standards for applicants, admitting all and only students who score at or above the standard, so might a conversational context set standards for 'tall'.[33] And maybe the same goes for 'knows'.

Keith DeRose often invokes the language of 'standards'. Here, for example, is a passage in which he characterises contextualism about 'knows' as

> a theory according to which the truth-conditions of knowledge-ascribing and knowledge-denying sentences (sentences of the form 'S knows that p' and 'S does not know that p' and related variants of such sentences) vary in certain ways according to the context in which they are uttered. *What so varies is the epistemic standard* that S must meet (or, in the case of a denial of knowledge, fail to meet) in order for such a statement to be true. In some contexts, 'S knows that *p*' requires for its truth that S have a true belief that *p* and also be in *a very strong epistemic position* with respect to *p*, while in other contexts, an assertion of the very same sentence may require for its truth, in addition to S's having a true belief that *p*, only that S *meet some lower epistemic standard*. DeRose (2009, pp. 2–3) (my emphasis)

I do not deny that talk like this can be useful. In fact, given my own preferred use of the relevant terms, which I'll outline below, a statement like this one from DeRose will come out as a true description of my preferred form of contextualism. But we should treat 'standards' talk cautiously, for at least two kinds of reasons.

First, as Stanley (2005, Ch. 2) has emphasised, knowledge ascriptions seem to behave in ways very different from gradable adjectives. For example, we don't seem to have any very natural way of talking about the relevant quantity that is suggested to come in degrees. 'Strength of epistemic position', unlike 'height', is a philosopher's term of art. We also don't have natural ways of expressing that one has *more* or *less* of the quantity in question, the way we so easily do with 'taller' and

know something about one's evidential state; the 'Speckled Hen Problem' shows that there are cases in which one's evidence doesn't even put one in a position to know certain features about one's evidence. (For the canonical presentation of the case and the related worry, see Chisholm (1942) and Sosa (2003).) Assume for the purpose of argument Lewis's phenomenological conception of evidence, and consider the experience one has upon looking at a hen with thirty-eight visible speckles; one's evidence entails that one has the visual experience as of thirty-eight speckles, but given ordinary discriminatory abilities, one won't know that one has the visual experience as of thirty-eight speckles (for all one knows, it's the experience as of thirty-seven speckles). One won't even know this if one somewhat bizarrely forms the rash belief that one has the experience as of thirty-eight speckles on the basis of one's experience. This isn't proper basing. Thanks to Baron Reed and Karen Lewis for bringing this issue into clearer focus for me. See Ichikawa and Jarvis (2013, pp. 310–1) for my own approach to the problem of the Speckled Hen, which precisely emphasises the distinction between sufficient evidence and (what I am here calling) proper basing.

[32] See e.g. (Cohen, 2005) and (Halliday, 2007).

[33] However, as Kompa (2002, p. 23) emphasises, even in cases of gradable adjectives, the relationship between contexts and standards can be exceedingly complex.

'shorter'. Relatedly, it is not a commitment of contextualism that the so-called 'epistemic standards' parameter be susceptible to a total ordering. The Lewisian view, for example, is most naturally developed in a way in which it can't. Suppose that Buttercup sees the lights on in the Captain's cabin; consider the sentence:

(5) Buttercup knows that the Captain is awake.

There are four contexts, c_1–c_4, according to Lewisian contextualism, for which each of these is true:

> c_1. (5) requires for its truth that Buttercup's evidence eliminate the possibility that Buttercup herself is asleep, merely dreaming that the Captain is awake with the lights on. (It also requires that more moderate skeptical scenarios, like those mentioned below, be ruled out.)
>
> c_2. (5) permits the proper ignoring of the dreaming hypothesis, but it does require eliminating certain specific skeptical possibilities, such as the possibility that the Captain might be sleeping somewhere other than his cabin.
>
> c_3. (5) permits the proper ignoring of the possibility that Captain is not in his cabin, but requires eliminating different specific skeptical possibilities, such as the possibility that the Captain has fallen asleep with the lights on. (It allows that the possibility that the Captain is sleeping elsewhere can be ignored.)
>
> c_4. (5) permits the proper ignoring of all possibilities in which the cabin light is on but the Captain is asleep. So 'knowledge' requires only eliminating the possibility that the lights are off.

If it isn't obvious that there are four contexts like this, stipulate that the light is a very good indicator of the Captain's wakefulness—good enough to make the existence of c_4 plausible. (This will be a rather non-skeptical context.) Then we can generate contexts with each of the other three features by adding salience of various kinds of error possibilities to the non-skeptical context.

Fig. 1.1 illustrates the point at issue. The entire diagram represents the space of all possibilities; the circles labelled with contexts indicate which worlds are relevant for the purpose of knowledge ascription (5). The dots correspond to particular skeptical scenarios—so the possibility that the Captain is sleeping elsewhere is among the relevant possibilities in c_1 and c_2; its being outside the c_3 and c_4 represents that it is properly ignored in those contexts.

There is a fine intuitive sense in which c_1 involves a higher epistemic standard than the others—it is a skeptical context, where a subject must rule out the possibility that she is dreaming, in order for a knowledge ascription to be true of her. And c_4 is very naturally understood as one that has a lower standard than the others. But these are all the comparisons that can straightforwardly be made. Notice that although c_2 and c_3 each require the consideration of certain slightly extraordinary possibilities, neither requires strictly more than the other. It does not seem right to say that either invokes a *higher* epistemic standard than the other: c_2 demands more in some respects; c_3 demands more in others. So if we want to talk about contexts supplying epistemic standards, we must be careful not to presuppose that these 'standards' can always be lined up in order.[34] We can describe a partial order well enough: standard s_1 is stronger than standard s_2 if and only if satisfying s_1 entails

[34] DeRose (1995, p. 31) countenances cases where 'it's unclear how the two epistemic positions we're evaluating compare with one another,' but seems to think that the unclarity at issue is merely epistemic. On the view I'm articulating, there are genuinely incommensurate strengths of epistemic standards; it is not merely a matter of our failure to recognise which is stronger.

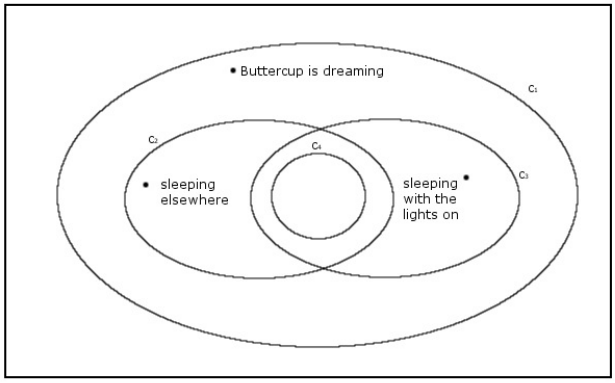

Figure 1.1: Four possible sets of relevant possibilities

satisfying s_2, but not vice versa. On the Lewisian framework, this will come approximately to the idea that the possibilities countenanced as relevant by s_1 are a superset of those that are relevant according to s_2. But there will be some incomparability.

This observation fits somewhat uncomfortably with Keith DeRose's treatment of epistemic standards, which characterises them in terms of reliability throughout modal space. In explaining his notion of a strength of epistemic position (to which epistemic standards are meant to correspond), DeRose writes:

> An important component of being in a strong epistemic position with respect to P is to have one's belief as to whether P is true match the fact of the matter as to whether P is true, not only in the actual world, but also at the worlds sufficiently close to the actual world. That is, one's belief should not only be true, but should be non-accidentally true, where this requires one's belief as to whether P is true to match the fact of the matter at nearby worlds. The further away one can get from the actual world, while still having it be the case that one's belief matches the fact at worlds that far away and closer, the stronger a position one is in with respect to P. DeRose (1995, p. 34)

According to DeRose, what it is to have a stronger epistemic position is to rule out the alternatives in broader sets of possible worlds. So what it is for knowledge ascriptions to require stronger epistemic positions is for them to require the ruling out of alternatives in broader sets of possible worlds.[35] An epistemic standard, then, could be thought of as a kind of measurement of modal distance—how far into modal space must one avoid error? If one draws a diagram of modal space, the radius of a circle centred on actuality could correspond to an epistemic standard. See Fig. 1.2. But if (as seems to follow from Lewisian contextualism) there can be contexts like c_2 and c_3 above, one cannot always represent the space of relevant possibilities as *spheres*; the possibilities corresponding to c_2 and c_3 have the same area and are both centred on actuality, but they are not

[35] See Blome-Tillmann (2008) for an argument against this way of explaining out standards. I find his argument somewhat compelling, although it probably turns on an open question about just how the modal similarity metric in question is supposed to be established.

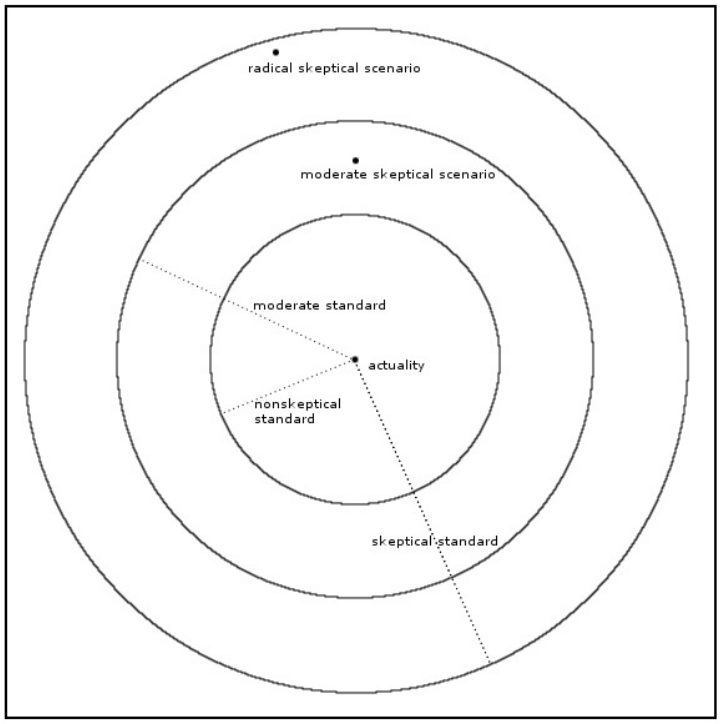

Figure 1.2: DeRose's model of epistemic standards as modal distances

congruent.[36] On DeRose's view, situations like the one described in Fig. 1.1, which involve overlapping sets of relevant alternatives, can never arise. One must ask which skeptical possibility is nearer in modal space: a case where the Captain is sleeping somewhere else, or one where he is sleeping in his cabin with the lights on? If one is relevant, then anything else closer is relevant too.[37]

I'll remain officially neutral in this book as to whether there can be a strict ordering on contexts of this kind, or whether the conversational pragmatics governing the set of relevant alternatives can take these kinds of less symmetrical shapes. Although I do think it would be a mistake to *assume* a strict ordering, I'll return to this question in Chapter 5, where we will see that there may be certain advantages to enforcing a more symmetrical approach.

[36] Schaffer (2005) also criticises DeRose's approach along these lines. Note however that he is using 'standards' in a stronger sense than I am—the idea that epistemic standards are what shifts is by definition, in Schaffer's framework, a matter of modal distance.

Compare also the discussion of 'epistemic standards' in Neta (2003), who argues that it is *that which counts as evidence*, rather than the epistemic standards, that shifts according to context—Neta's assumption seems to be that the 'standard' is or supervenes on the strength of evidence that is required for knowledge. On the more general notion of 'standards' I am using, an approach like Neta's according to which contextual variation makes for a difference in what is 'known' because it makes for a difference in what is 'evidence' will count as one in which the standards change.

[37] Perhaps one might hold on this view that we don't typically know which is a nearer skeptical possibility. Then if one is relevant, we may not know whether the other is relevant. In such cases, if Buttercup's evidence eliminates one skeptical possibility but not the other, we wouldn't be able to know whether, in certain contexts, the knowledge ascription could be true.

The second kind of reason we must be careful about 'standards' language is that conversational context is not the only thing that can influence which alternatives are relevant. Invariantists sometimes talk about varying 'standards' too. Getting clear on what this suggestion amounts to is crucial for distinguishing contextualism with rival views like 'interest-relative invariantism'. I turn to these issues now.

1.9 Invariantism with Shifting Standards

Consider, for example, 'interest-relative invariantism,' (IRI) the family of views that has recently been defended by Stanley (2005), Fantl and McGrath (2009), and Weatherson (2011).[38] According to these views, some practical considerations, such as how important a question is to a subject, play roles in determining what a subject must be able to rule out in order to have knowledge.[39] IRI is often described as similar to contextualism in that differing 'standards' govern the truth of a knowledge ascription, with the difference being that contextualists think the *speaker's* interests are relevant for determining the standards, while IRI has it that the *subject's* interests are what matter.

Features other than practical interests are also sometimes thought, by other kinds of 'shifty invariantists', to play such roles. For example Hawthorne (2004) suggests that which possibilities are *salient* to a subject make a difference for what one has to do in order to know; Wright (2011) says that features of the subjects' *social roles* make a difference.[40] So if the epistemic standard is characterised in terms of which possibilities must be ruled out, then shifty invariantism, like contextualism, will posit shifts in epistemic standards. DeRose identifies and accepts this consequence of his view, writing that shifty invariantism 'agrees with contextualism that varying epistemic standards govern whether a speaker can truthfully claim "I know that p".'[41]

It is worth appreciating, however, that if epistemic standards are characterised in this way, then it is not only contextualists and shifty invariantists who are committed to varying epistemic standards for knowledge or knowledge ascriptions. Probably every nonskeptic is committed to *this* kind of 'shifting standards'. Certainly, anyone who embraces orthodoxy about Gettier cases must also think that which possibilities a subject must rule out in order to satisfy 'knows' can vary according to the subject's situation.[42] For example, suppose that Ralph believes on the excellent grounds that Josephine's social class is above his own. Suppose also that Ralph loves Josephine, and knows that he does. Naturally, Ralph believes that he loves someone of a different class. If all were as it seems, then this belief would amount to knowledge; but in fact, Ralph has been misinformed about each of their social positions—he himself comes from a noble family, while Josephine's roots are

[38] This family of views is sometimes labelled 'subject-sensitive invariantism' or 'SSI'. This is not a particularly apt name, as it is trivial that whether a subject has knowledge depends in part on features of the subject. Cf. Stanley (2005, p. 122). My preferred 'interest-relative invariantism' makes explicit that knowledge depends on facts about the subject's *interests*, not just about the subject generally. It should be noted, however, that this name has its shortcomings too; the view of Hawthorne (2004) for instance, is typically categorised alongside the others mentioned in the main text, even though Hawthorne's suggestion is that knowledge depends upon what is *salient* to the subject; it is something of a stretch to characterise this as a fact about the subject's *practical interests*. I use 'shifty invariantism' as a more general term. But in light of what I will go on to say about epistemic standards, interest-relative invariantism, and Gettier cases, articulating what it is that characterises shifty invariantism in general is a surprisingly difficult question.

[39] Because these views are invariantist, it is more natural to speak at the object level about what is required for *knowledge*, rather than for the truth of a *knowledge ascription*. Strictly speaking, however, one could put things into the metalinguistic terms: according to IRI, whether a knowledge ascription is true of a subject depends in part on the subject's practical interests. The IRI theorist simply disquotes: in general, if invariantism of any kind (including IRI) is correct, a knowledge ascription is true of a subject iff the subject has knowledge.

[40] The 'methodological contextualism' of Wright (2010), which she derives from Williams (1991) and Williams (2004), is also a shifty invariantist view of this kind.

[41] DeRose (2009, p. 108)

[42] See Brueckner (1994) on this point.

working-class. When he believes that he loves someone of a different class, his belief is true, but it is not knowledge; Ralph is in a Gettier case. According to invariantist orthodoxy, Ralph's epistemic position is sufficient for knowledge in the case where everyone's social status is as it appears, but not in the case where it turns out that the evidence about social status was misleading. So when the environment is less cooperative, it appears that the 'epistemic standards' (in DeRose's sense) necessary for knowledge are higher. This is a conclusion that I doubt DeRose would want to accept, as he characterises contextualism and shifty invariantism as *distinctive* in allowing for differing standards governing knowledge ascriptions.[43] But it is not clear how to articulate the notion of an epistemic standard in a way that makes the mechanisms of IRI views amount to changes in standard, without categorising Gettier cases in the same way.

Indeed, one needn't appeal to Gettier cases to make this point—once the thought is articulated clearly and laid out for examination, it's just *obvious* that different possibilities can be relevant, depending on the subject's situation. All theorists will accept, for instance, that whether one needs to rule out the hypothesis that the animal before one is a cleverly disguised mule depends at least in part on whether there's any reason to think it might be.

Epistemic standards are sometimes glossed as matters of what *purely epistemic* factors are required for knowledge or knowledge ascriptions;[44] they are also sometimes described as the *truth-relevant* factors that are required.[45] But the notions of 'purely epistemic' factors and 'truth-relevant' factors are inevitably left unexplicated; I myself have never been confident that I have a grip on what they amount to.[46] There is perhaps a response-dependent property in the neighbourhood: one might identify 'epistemic standards' as governing only those features that *obviously* or *unsurprisingly* influence knowledge or knowledge ascriptions. But there is no clear role for such response-dependent properties in serious theorising about knowledge. So I don't find it useful to categorise contextualism and shifty invariantism alike as involving the use of different 'epistemic standards'.

Notice also that it should be possible to endorse both contextualism and e.g. the kind of interest-sensitivity characteristic of IRI. The 'invariantism' in 'interest-relative invariantism' is the negation of contextualism, so one can't coherently be a contextualist who endorses IRI, but the semantic thesis of invariantism about 'knows' is perfectly consistent with the metaphysical thesis of 'pragmatic encroachment' posited by IRI[47]—namely, that knowledge depends on the practical situation of the subject.[48] But if such pragmatic encroachment is characterised as the view that the epistemic standards are fixed by the practical situation of the subject, then one cannot consistently hold contextualism, if it is characterised as the view that the speaker's context fixes the standards. The same goes, mutatis mutandis, for other kinds of shifty invariantism, such as the social-sensitivity view of Wright (2011).

One *could* describe the 'interest-relative contextualist' view I am mentioning as a view according to which the speaker's context and the subject's practical situation *both* play roles in fixing the

[43] I don't have the space to make the case here, but my view is that the arguments from e.g. (DeRose, 2009, p. 189) turn on this commitment.

[44] Fantl and McGrath (2009, p. 35) describe 'impurism' as the thesis that '[h]ow strong your epistemic position must be—which purely epistemic standards you must meet—in order for a knowledge-attributing sentence, with a fixed content in a fixed context of use, to be true of you varies with your circumstances.' DeRose (2009, p. 36) likewise characterises standards as a matter of 'how good an epistemic position' one needs to know.

[45] E.g. Stanley (2005, p. 1), Fantl and McGrath (2009, p. 557–8), Grimm (2011), DeRose (2009, pp. 24 & 185), Brown (2012b, p. 43).

[46] If, as one might've found pre-theoretically plausible (obvious?), knowledge itself is both 'purely epistemic' and 'truth-relevant', then it is trivial that a fixed epistemic standard—viz., knowledge—is necessary and sufficient for knowledge, in all cases. So whether or not one adopts shifty invariantism, one is committed to the idea that there is a purely epistemic, truth-relevant property that is necessary and sufficient for knowledge. If knowledge *isn't* 'purely epistemic', or if it isn't 'truth-relevant', then I don't know how to determine when those terms apply.

[47] The term 'pragmatic encroachment' is due to Jonathan Kvanvig in a 'Certain Doubts' blog post in June 2004.

[48] Cf. McGrath (2010) and Fantl and Mcgrath (2012) who also point out this consistency.

epistemic standards necessary for the truth of a knowledge ascription, but this obscures the fundamental difference between contextualism—a semantic thesis about the verb 'knows'—and pragmatic encroachment—a metaphysical thesis about knowledge. It better gets at what is distinctive of these views, I suggest, to describe contextualism as the view that different standards govern knowledge ascriptions in different contexts, while pragmatic encroachment holds that practical considerations about a subject are relevant to determining whether that subject meets a given standard. Consequently, I do *not* draw the straightforward connection between epistemic standards and contrast classes. According to orthodoxy, whether there are fake barns around influences which worlds are relevant; according to pragmatic encroachment theorists, the practical importance of a question to the subject also influences which worlds are relevant. I do not describe either kind of subject-sensitivity as a shift in the *standards* that must be met for knowledge.

I recognise this as a terminological departure from DeRose, whose seminal role in setting out this terrain admittedly entitles him to a degree of stipulative authority over the vocabulary in the area. Nevertheless, I consider this departure sufficiently theoretically motivated to justify my slightly heterodox use of terminology here. My view is that epistemic standards are simply functions from indexes to contrast classes.[49] One can recover the idea of a stronger or weaker standard thus: standard A is stronger than standard B (and B is weaker than A) iff, for every index i, A(i) is a proper superset of B(i). This will have the implication that the satisfaction of standard A will entail the satisfaction of standard B. Since I agree with DeRose that contextualism is well-characterised as the thesis that knowledge ascriptions require the satisfaction of different standards in different contexts, I also agree with DeRose that contextualism allows for different alternatives to be relevant in different contexts. In fact, my approach is easily read as a generalisation on DeRose's remarks here: he thinks that standards interact with indexes to generate sets of relevant worlds by drawing a sphere with a particular radius around them in modal space; I think standards interact with indexes to generate sets of relevant worlds in a more complicated way. But unlike DeRose, I do not consider shifty invariantism to be like contextualism in allowing different standards to govern different knowledge ascriptions about different situations. If invariantism is true, there is only one epistemic standard; so *for any given index*, there is only one set of relevant alternatives. But the index encodes information about the practical situation of the subject, which may be relevant for the question whether a given standard is met.

1.10 Factivity

Something DeRose and I do agree about, however, that it won't do to identify standards with sets of relevant possibilities.[50] One problem with this idea is that, if standards are supposed to be set by contexts, it asks contexts to do work that indexes should be doing. The easiest way to see this, though not the only way, is by considering factivity. Whether a knowledge ascription is true depends in part on whether the content of that ascription is true at the index at which the knowledge ascription is evaluated. That is to say, in the Lewisian framework, the world of the index is always relevant, for any knowledge ascription.[51] But it is not very plausible that in general, the conversational context will supply the needed facts about the subject's situation.

To see this, suppose that Dick has heard Ralph discussing a plan to elope, and that Buttercup knows that Dick has done so. On this basis, it is plausible to suppose that there is a context in

[49] DeRose also thinks of standards as such functions—see DeRose (2009, pp. 37–8). But as explained in the main text, his view—at least the version in DeRose (1995)—has it that the functions have rather different characters to the ones I have in mind.

[50] See DeRose (2009, pp. 35–41).

[51] Cf. Lewis (1996, pp. 554–5); see also my §1.7.

which Buttercup could truly say, 'Dick knows that Ralph is planning to elope.' But consider the same sentence, in the same context, in a world where things seem the same to both Dick and Buttercup, but where Ralph is *not* planning to elope. (Perhaps his discussion of such putative plans was intended to deceive would-be meddlers like Dick.) In this world Ralph is not planning to elope; so factivity demands the relevance, for Buttercup's knowledge ascription, of the possibility that Ralph will not elope. Since this possibility is a case where Ralph has no plans to elope, and is uneliminated by Dick's evidence, Buttercup's knowledge ascription in this context must be false. Schematically, where C is the nonskeptical context in question, S is the sentence, 'Dick knows that Ralph is planning to elope', w_1 is the world where things are as they seem, w_2 is the world where Ralph's conversation was misleading, and X is the skeptical possibility that Ralph will not elope even though the evidence is as it is in w_1:

1. X is a possibility in which Ralph is not planning to elope, and it is uneliminated by Dick's evidence.
2. In w_1, Buttercup's utterance of S in C is true.
3. So in w_1, X is an irrelevant possibility.
4. In w_2, X is the case. This explains why in w_2, Buttercup's utterance of S in C is false; so X is relevant.
5. So C does not fix whether X is relevant.

If this argument is right, then it would be a mistake to hold a version of Lewisian contextualism according to which the context fixes the relevant alternatives; if context is supposed to fix epistemic standards, then epistemic standards cannot be identified with—or even determine—the sets of relevant alternatives. Insofar as one wishes to characterise a plausible version of Lewisian contextualism, then, one should not interpret the view as Jonathan Schaffer does, according to which '[a] sentence of the form "s knows that *p*" is true in context c iff s's evidence eliminates every not-p possibility relevant in c.' Schaffer (2015, p. 475)[52]

There is, however, a substantive assumption in the argument given above. As discussion of this point is somewhat technical and might be considered peripheral to readers not invested in the debate, I'll discuss it in its own section; some readers may prefer to skip ahead to §1.12.

1.11 Modality and Knowledge Ascriptions

The argument of the previous section assumed that the very same context, C, might occur in the relevantly distinct possible worlds w_1 and w_2. The intuitive thought underwriting this assumption was that the difference between w_1 and w_2 was in some sense wholly external to the experience

[52] Another reason it's important to be precise about how we think about standards, and just what is fixed by context, is that failure to do can make contextualism appear susceptible to objections it needn't be vulnerable to. For example, Jason Stanley argues that contextualists can't comfortably allow that subjects' practical situations influence how easy it is to satisfy 'know', in cases where the speakers are unaware of the practical importance to the subject ('Ignorant High Stakes' cases). Stanley writes that for contextualism to give the right result, 'it must turn out that the fact that there is a greater cost to Hannah's being wrong affects the semantic content of some of her statements, even when neither she, nor any other conversational participant, is aware of it. So the contextualist could accommodate Ignorant High Stakes, but only at the cost of advancing a rather dramatic claim about the potential semantic effects of non-psychological facts about extralinguistic content.' Stanley (2005, p. 26). Even granting for the purpose of argument that such an implication is a costly one, the contextualist can avoid it by clarifying that the 'standard' provided by features of the context of which the speaker is cognisant itself encodes a sensitivity to the subject's practical situation.

of Buttercup, the speaker; whether Ralph has been secretly deceptive or not isn't relevant to what kind of context she is speaking in. (Ralph is not a participant in the conversation.) But in fact, this was a stronger assumption than I needed. The key assumption is that Buttercup expresses the *same proposition* in w_1 as she does in w_2. (This is weaker than the assumption that the contexts are identical, because some differences in conversational context can be idle for the purpose of a given knowledge ascription. So what is really at issue here is whether the two contexts are the same *in all relevant respects*—i.e., whether the knowledge ascription expresses the same proposition in each of them.) The argument assumes that that *what Buttercup says* does not depend on whether Ralph's conversation was sincere. Whether what she says is *true* depends on Ralph, but not what she says itself.

It should be admitted, however, that these kinds of higher-order metasemantic intuitions are rather theoretical; someone might deny the intuitions cited above about whether Buttercup expresses the same proposition in distinct possible worlds which she cannot distinguish. I think this would be a counterintuitive move, but not a disastrous one. Jonathan Schaffer has told me in conversation that this is his preferred way to think of things. So, adapting his version of the view to my example, Schaffer would insist that the contexts mentioned in (2) and (4) above are in important ways distinct: the context mentioned in (2) is one in which Buttercup is discussing Dick's position vis-à-vis Ralph *in* w_1; the context mentioned in (4) is one in which Buttercup is discussing Dick's position *in* w_2. This difference in context does make for a change in the relevant alternatives. So what is said *is* different between w_1 and w_2. But there is another reason not to go this way, in addition to its counterintuitive consequences about what is said. Taking the Schaffer route commits one to choosing between implausible verdicts about modal profiles of propositions expressed by knowledge ascriptions.

Consider again Buttercup's knowledge ascription in distinct worlds w_1 and w_2. In w_1 it is true, and in w_2 it is false; Schaffer thinks that this is because she expresses different propositions—the sentence expresses p_1 in w_1, and p_2 in w_2, where p_2 is stronger than p_1 in that it entails elimination of the possibility that Ralph was lying. Consider p_1, the true proposition Buttercup expressed with 'Dick knows that Ralph is planning to elope' in w_1. Its truth requires Dick to rule out cases where Ralph is (a) not planning to elope, and (b) straightforward and honest in his reports about his intentions. Dick can do this in w_1, since he heard Ralph say that he was planning to elope. That's why p_1 is true in w_1. But notice that p_1 *also* looks to be true in w_2, where Ralph was lying about his plans to elope. For p_1 does not require Dick to rule out the possibility that Ralph was lying. On the Schaffer model, where epistemic standards are fixed by contexts, and identify sets of relevant alternatives, we must say that the proposition expressed in w_1 is one that doesn't imply elimination of that skeptical possibility. But this means that, relative to a world where that possibility obtains—in this case, w_2—that proposition may still be true, even though that said to be 'known' is false. That is to say, what Buttercup says with her utterance of 'Dick knows that Ralph is planning to elope' in w_1—p_1—is true in some worlds where Ralph isn't planning to elope, viz., w_2. This is to nothing less than a violation of factivity in the straightforward sense: it is false, on the view I am here sketching, that the proposition expressed by 'S knows that q' entails that q. There are possible worlds—w_2 in my example—where the proposition is true, but q is false. This by itself is a serious theoretical cost; it also has particular counterintuitive implications. For example, when Buttercup expresses p_1 in w_1, it is extremely natural to suppose that if Ralph had been lying, what Buttercup said would have been false. But what Buttercup said, on the view we're considering, was p_1; and p_1 is true at worlds like w_2 where Ralph was lying.

The considerations just raised should put pressure on most contextualists away from thinking of contexts as establishing sets of relevant alternatives, the way that Schaffer does. However, I do not expect these considerations to have much sway against Schaffer himself. For Schaffer is independently committed to *necessitarianism about propositions*—the view that all propositions are

ultimately propositions about particular possible worlds, and so have their truth values necessarily. (Schaffer, 2012). So Schaffer is already committed to the view that Buttercup's sentence in w_1, 'Dick knows that Ralph is planning to elope', expresses a proposition that is true in w_2 where Ralph is not planning to elope. On Schaffer's view, the proposition in question is something like the proposition that *in w_1, Dick knows that Ralph is planning to elope, rather than sitting around and not doing or saying anything about eloping*; and this proposition is necessarily true. Even relative to worlds in which Dick has never heard of Ralph, it's still true that *in w_1*, Dick knows that Ralph is planning to elope. Whether this amounts to a violation of factivity turns on sensitive questions about what exactly factivity requires—it *is* a view according to which the proposition expressed by 'S knows that p' does not *entail* that p, in the usual sense that there are possible worlds where the former proposition in question is true but the latter is false. But if it is a failure of factivity, it is one that the necessitarianism is already committed to.

Insofar as one wishes not to commit the Lewisian contextualist to necessitarianism about propositions, one has good reason not to think that context fixes the relevant alternatives.

1.12 Differences between Knowledge and Quantifiers

Jason Stanley has argued that knowledge ascriptions interact with contexts in a way quite different from the way standard quantifiers like 'every' and 'some' do. According to Stanley, this undermines the Lewisian strategy of understanding knowledge as context-sensitive along the model of quantifiers.[53] Stanley defends the following claim about context-sensitive discourse in general:

> Since semantic context-sensitivity is traceable to an individual element, multiple occurrences of that element in a discourse should be able to take on differing values. In the case of an utterance such as 'This is larger than this', where two different objects are pointed to by the person uttering the sentence, this feature is obviously confirmed. But it is present in a broader range of constructions. Stanley (2005, p. 57)

Call the ability of context-sensitive terms to take on different values rather freely *flexibility*. Stanley suggests that all context-sensitive discourse is flexible,[54] and that if 'knows' is likewise flexible, then contextualism about 'knows' does not enjoy the advantages claimed for it. For example, Stanley writes,

> If different occurrences of instances of 'knows that *p*' can be associated with different epistemic standards within a discourse, some of the paradigm sentences the infelicity of which supposedly motivates [contextualists'] accounts over rival accounts turn out to be felicitous and potentially true by contextualist lights. For example, if we have similar behavior to many other context-sensitive expressions, one would expect the following to be felicitious:
>
> (6) Bill knows that he has hands, but Bill does not know that he is not a bodiless brain in a vat. Stanley (2005, p. 67) (my numbering)

[53] In fact, Stanley's argument is intended to generalise to all forms of contextualism, but it is the Lewisian strand that interests me here.

[54] One might perceive an obvious potential objection to this claim in, to use Kaplan's term, 'pure indexicals'—indexicals like 'I' that take their semantic value automatically, without the aid of any kind of demonstrations. See Stanley's pp. 65–6 for a rejoinder.

There are two challenges for the contextualist here. First, a major historical motivation for contextualism has been its supposed ability to resolve skeptical puzzles without admitting the truth of sentences like Stanley's (6); if Stanley is right that contextualism predicts (6) to have easy true readings, a motivation for contextualism is undermined. Second, more straightforwardly, (6) is intuitively repugnant; any view that implies it should be rejected.[55]

In support of his contention that Lewisian contextualism predicts sentences like (6) to be unproblematic, Stanley cites sentences involving explicit quantifiers whose domains shift mid-sentence. For example, Stanley points out that it's not difficult to find a reading of 7 according to which the two instances of 'every' take different domains:

(7) Every sailor waved to every sailor,

For example, if two ships have recently crossed paths, it might say that every sailor on one ship waved to every sailor on the other. More directly to the point, Stanley also offers this dialogue:

A. Every van Gogh painting is in the Dutch National Museum.

B. That's a change; when I visited last year, I saw every van Gogh painting, and some of them were definitely missing. Stanley (2005, p. 65)

Stanley suggests that in the natural reading of this dialogue, the 'every' in (A) and the 'some' in (B) each range over all van Gogh paintings in existence, but that the 'every' in (B) is restricted to those paintings that were in the museum at the time of the visit. I agree. So Stanley is right that the domain of 'every' can become larger or smaller over a relatively short period within a conversation. But this does not imply, as Stanley suggests it does, that Lewisian contextualism predicts abominable conjunctions to be acceptable. It does not even imply that there is any significant disanalogy between knowledge ascriptions and quantifiers. Stanley has shown that 'every' is at least somewhat flexible, but he has not shown that it is as flexible as the disanalogy would demand.[56]

Consider this sentence:

(8) Bill's evidence eliminates all possibilities in which he lacks hands, but Bill's evidence does not eliminate all possibilities in which he is a handless brain in a vat.

This sentence is closely connected to the abominable conjunction (6) mentioned above, given Lewisian contextualism. But it is not at all easy to find a true reading of (8). In fact, (8) feels like a contradiction, much as the simpler (9) does:

(9) Every sailor is on deck, but some sailors are below deck.

But (8) and (9) are not obviously coherent; one cannot easily find true readings of them. This is a point of analogy, not of disanalogy, between quantifiers and 'knows'; for (6) works just the same way.[57]

So it does not seem to me that contextualists should accept that there is enough flexibility in 'knows' to predict the felicity of abominable conjunctions. This is not to insist, however, that 'knows' is totally inflexible, either—and the same goes for the case of quantifiers. For example, if

[55] On both of these points, see especially DeRose (1995).

[56] I first gave this argument in defense of Lewis against Stanley's attack in Ichikawa (2011a, pp. 391–92). For a similar response, see Blome-Tillmann (2014, pp. 87–94).

[57] Gauker (2010) argues that Stanley's own semantics for quantifiers (Stanley and Szabó, 2000) implies that there are true readings of (8) and (9), and uses this fact to motivate his own alternative. He also points out (p. 267) that 'knows' patterns with quantifiers in this way.

one adds supporting context to help make clearer the apparent shift in epistemic standards, it might be possible to get something like the mid-sentence shifts Stanley describes. I think (10) is much preferable to (6), for instance:

> (10) Since he certainly wouldn't fail to notice an amputation, Bill knows he has hands, but since he can't prove it from first principles, Bill doesn't know he's not a disembodied brain in a vat.

Or, sticking to more mundane skeptical scenarios:

> (11) Having already met the family, Joseph knows that the Captain has a daughter, but being susceptible to practical jokes, Joseph doesn't know that he's not the victim of an elaborate conspiracy about the captain's family.

I certainly do not think that (10) and (11) are obviously fine, but they strike me as considerably less abominable than the original (6). It may be that 'knows' is flexible enough to accommodate true readings of sentences like these; if so, a Lewisian contextualist needn't quarrel with Stanley's claim that all context-sensitive terms are (to some degree) flexible. So *contra* Stanley, it does not seem to me that there is any strong disanalogy between knowledge ascriptions and quantifiers with respect to flexibility.

1.13 Knowledge Embedded in Conditionals

So I'm unconvinced by Stanley's argument that knowledge ascriptions behave importantly differently from quantified sentences. On the contrary, I think there are deep *similarities* that haven't been sufficiently appreciated. In this section I'll exploit one of them to give a positive linguistic consideration in favour of contextualism. So far as I can tell, this point has not been made in the literature before.

As noted in §1.3 above, modals are context-sensitive. Words like 'must,' 'may,' 'probably,' etc. are quantifiers over possibilities, and their domains are influenced by the context of utterance. Modals also interact in a characteristic way with conditionals. Bob knows that Joseph and Hebe always travel together, but he doesn't know whether they've arrived yet. Since he doesn't know whether either has arrived, (12) expresses a truth in Bob's context:

> (12) Hebe may not have arrived yet.

The 'may' here is a possibility modal, quantifying over the worlds consistent with Bob's knowledge.[58] So (12) says that, of the relevant worlds, there are some in which Hebe has not yet arrived. The 'must' in (13) is also epistemic:

> (13) If Joseph has arrived, then Hebe must have arrived.

Since Bob knows that Joseph and Hebe invariably travel together, (13) also expresses a truth in Bob's context. If we also stipulate that Joseph has in fact arrived (although Bob doesn't know it), then (14) also obviously expresses a truth in Bob's context:

> (14) Joseph has arrived.

[58] I am glossing over some subtleties about epistemic modals here. For example, I ignore whether Bob's knowledge alone is relevant, or that of his conversational community. I also ignore whether it is Bob's *knowledge* that is relevant, or something else—perhaps his evidence, if evidence is distinct from knowledge. (Though as I'll suggest in Ch. 3, I think it's not.)

Admittedly we wouldn't expect Bob to *utter* (14), since he doesn't know the proposition it expresses. But it still expresses a truth in his context. The content of (14), unlike that of (12) or (13), has nothing to do with Bob's epistemic position.

Notice that this would make trouble given a principle like this:

> Naive Modus Ponens: For any context C, if, for some independent clauses 'A' and 'B', the sentences 'if A, then B' and 'A' each express a truth in C, then the sentence 'B' expresses a truth in C.

Naive Modus Ponens, combined with the observations just made about (12)–(14), entail that 'Hebe must have arrived' is true in Bob's context. But this is inconsistent with (12), which is true. So Naive Modus Ponens is false in generality.[59] The explanation for its particular failure here is that contrary to appearances, the deeper form of (13) is not a propositional connective relating the proposition expressed by 'Joseph has arrived' to that expressed by 'Hebe must have arrived'. Rather, the effect of the 'if' clause is to *restrict* the modal 'must'.[60] Ignoring any possibility where Joseph hasn't arrived, Hebe must have arrived.

This is all very different from the way that conditionals interact with non-modal claims. These really are inconsistent:

(15) If Joseph has arrived, then Hebe has arrived.

(16) Joseph has arrived.

(17) Hebe has not arrived.

Naive Modus Ponens gets the right result here. It is characteristic of the context-sensitive modals that they interact with conditionals in a way that invalidates Naive Modus Ponens. And tellingly, knowledge patterns with the modals. Suppose as before that Bob doesn't know whether Joseph has arrived. He also doesn't know what kind of mood the Captain is in. Still, Tom might truly utter to Bob:

(18) If Joseph has arrived, you know the Captain is nervous.

This might be true even though (19) and (20) are both true in Tom's context:

(19) Joseph has arrived.

(20) Bob doesn't know that the Captain is nervous.

Although such sentences are not difficult to find in the wild, philosophers who've trained themselves to think in formal logic sometimes need a little help hearing the intended reading of (18). I think that (18) is ambiguous; there is a less interesting reading, for my purposes—one philosophers are likelier than others to hear—where it says that in all relevant worlds where Joseph has arrived, Bob knows that the Captain is nervous. But focus on the reading that is tantamount to Bob's knowledge of the conditional, *if Joseph has arrived, then the Captain is nervous*. (Compare: 'Don't even bother asking him; after all, if it's true, you know he'll just deny it anyway'.)

[59] Compare the 'odious inferences' puzzle of Dowell (2012, pp. 272–4); as Dowell indicates, a kind of 'dynamic modus ponens' is plausibly valid; see her pp. 285–88.

[60] I am assuming something like the framework of Kratzer (2012). For a helpfully accessible overview, see Dowell (2012, pp. 283–5), Portner (2009, pp. 81–2), or Bronfman and Dowell (2016).

Naive modus ponens goes wrong here. Joseph has arrived; if Joseph has arrived, then Bob knows the Captain is nervous; but Bob doesn't know that the Captain is nervous. The Lewisian contextualist has an appealing explanation for this pattern: knowledge is a kind of epistemic modal, and like modals generally, it can have its scope restricted by if-clauses. A sentence of the form 'if A, S knows *p*' says, first restrict your attention only to the A worlds; S's evidence eliminates the not-*p* worlds that remain.

1.14 Is Contextualism Ad Hoc?

Before closing this chapter, I'd like to return to the broader questions about the theoretical significance of knowledge, setting the stage for much of the rest of the book. Let's assume that some kind of relevant alternatives approach captures the truth conditions of 'knows' ascriptions. For a subject to satisfy 'knows *p*', the subject must rule out all relevant alternatives to *p*. According to my brand of Lewisian contextualism, which possibilities are among the relevant alternatives depends on features of both context and index. One of the chief virtues of contextualism is that, compared with its competitors, it allows us to preserve more of our intuitions about truth conditions. When we are worrying about skeptical scenarios, we are inclined to say that someone 'doesn't know that *p*'; when we aren't, we are inclined to say someone 'does know that *p*'. Where invariantists offer an error theory here, the contextualist can have it that both inclinations are correct.[61]

In making this point so far, I have focused on skeptical intuitions arising from radical skeptical scenarios. But there is a wide diversity of cases that elicit contextualist-friendly patterns. For example, 'bank cases', made famous by DeRose (1992, p. 913) seem to show that when questions are more practically important, we seem to require more of subjects in order to attribute 'knowledge'. Contextualists accommodate these intuitions straightforwardly—in contexts in which the question is important, or in which possibilities of error are foregrounded, higher epistemic standards are in play.

But at what cost comes this capacity to accommodate intuitive data? One might worry that, given the ease with which our intuitions are manipulated, any theory that accommodates all of them would ultimately be ad hoc. Contextualism performs well at capturing the intuitive truth conditions of knowledge ascriptions, but maybe those truth conditions are an unholy mess. Does capturing them require an ad hoc semantics? In engaging in semantics, capturing intuitions is one virtue, but theoretical significance and systematicity are another.[62] Do the intuitive benefits justify the added theoretical complexity?

This is a worry that should be taken seriously. However, it is broader than it first appears, in at least two respects. First, it is not specific to relevant alternatives approaches to contextualism. *Any* approach that uses distinct standards for knowledge ascriptions in different contexts faces the question of whether the cited standards correspond to an interesting category, or whether they are merely an ad hoc device to preserve our intuitions. For example, a contextualist who thinks that the *degree of justification* necessary for knowledge ascriptions varies with context will owe an explanation of the value of the decreased systematicity.

Second, it is not only contextualists who face this challenge; invariantists also owe an answer to a version of it.[63] An invariantist holds that there is a single privileged epistemic standard that is in play

[61] As I'm using the term, an 'error theory' is simply a view attributing error to ordinary speakers—invariantists will say either that our intuitions in skeptical cases, or in non-skeptical contexts, are mistaken. But the error theory doesn't imply that knowledge discourse isn't often correct. As Nate Bemis reminded me, in some philosophical contexts the term 'error theory' is used for the much stronger claim that a whole domain of sentences are uniformly false—see the discussion in Joyce (2015, §4). No such general claim is here intended.

[62] See e.g. Haslanger (2006) and Weatherson (2003).

[63] Relativists do too, very similarly to the way contextualists do.

for all knowledge ascriptions. But what is special about this privileged standard? Consider, from the point of view of an invariantist, a case of 'barely-knowledge' and one of 'almost-knowledge'—maybe the one and only standard for knowledge is barely met by my belief about the current location of my car, but barely missed by my belief about what I'm going to cover in class next Friday.[64] It also seems at least potentially ad hoc for the invariantist to draw the line at a particular point. One might wonder what theoretical justification there is for highlighting that particular standard as knowledge-defining.[65]

The way to engage with these questions, I suggest, is to investigate the theoretical roles that the standards in question will play. If we find that there does seem to be a metaphysically privileged standard in the neighbourhood of how we intuitively categorise knowledge ascriptions, this will be suggestive of the invariantist approach. On the other hand, if we find that the epistemic standards parameter invoked by the contextualist has theoretically significant roles to play *beyond* explaining our intuitions about knowledge, then the contextualist invokes phenomena in which we already have independent reason to believe. Such is the main strategy of the remainder of this book. The context-sensitivity of knowledge ascriptions, I will suggest, can be explained through variation in epistemic standards, where that variation itself does independent work. In particular, such variation is helpful in understanding the broader theoretical roles of knowledge. These epistemic standards that I posit in my approach to knowledge have general theoretical work to do; understanding them will help us to understand the theoretical significance of knowledge.[66] This is how contextualism and the knowledge first project can mutually support one another, and this is my aim in this book.

[64] Just how this would be motivated would depend on the details of the invariantist approach, but one might think, for example, that a certain degree of *safety* is required, and that possibilities where e.g. my car has surprisingly been towed since this morning are in the relevant sense *slightly* more distant than ones where e.g. I decide after Wednesday's class that we really ought to spend one more day talking about hermeneutical injustice. If so, I may know where my car is, but fail to know what we'll do in class on Friday. An invariantist might quibble with any particular such pair, saying I know both or neither, but any nonskeptical invariantist must think there are some pairs of cases like this.

[65] This is a version of a challenge laid out in BonJour (2010, pp. 39, 42), who concludes in part on these grounds that knowledge is of limited epistemic interest.

[66] Compare Lewis's famous remarks in defence of his posit of the existence of non-actual possible worlds:

> Why believe in a plurality of worlds?—Because the hypothesis is serviceable, and that is a reason to think that it is true. The familiar analysis of necessity as truth in all possible worlds was only the beginning. In the last two decades philosophers have offered a great many more analyses that make reference to possible worlds, or to possible individuals that inhabit possible worlds. I find that record most impressive. I think it is clear that talk of *possibilia* has clarified questions in many parts of the philosophy of logic, of mind, of language, and of science—not to mention metaphysics itself. Lewis (1986, p. 3)

Chapter 2

Sensitivity

One of the more intuitive responses to the Gettier problem was the invocation of a connection between knowledge and certain counterfactual conditionals.[1] Suppose that Chloe, a student at a girls' school, believes on strong but misleading evidence that a certain student is a man disguised as a woman. From this belief she draws the straightforward conclusion that there is a man at the school. The student in question is in fact not a man—but as it happens, there is a man who has just climbed over the fence. Chloe is like the subjects of Gettier's famous paper; she justifiably believes that there is a man at the school, but she doesn't know it.[2] A very natural explanation for her failure to know is that her belief is *insensitive* to the truth; that is, *even if there hadn't been a man at the school, she still would have believed that there was*. A generalisation of this explanation has it that such sensitivity is a necessary condition for knowledge.[3]

Since its late-twentieth-century heyday, the idea that sensitivity is necessary for knowledge has fallen out of favour. I perceive three principle motivations for its ebb of popularity. First and most notoriously, the sensitivity theory appears to licence unacceptable violations of closure; my belief that I have hands is sensitive, and so plausibly knowledge, but my belief that I'm not a brain in a vat is insensitive, and so cannot be knowledge. But the conjunction—Jonathan knows he has hands, but does not know that he's not a brain in a vat—is abominable. Second, there appeared to be a class of counterexamples to sensitivity—cases of insensitive belief that were nevertheless intuitively cases of knowledge.[4] These cases suggested that we might prefer a different modal relation between belief and truth as necessary for knowledge. Third, sensitivity did not, on further examination, seem to solve the Gettier problem in generality, as there were Gettier cases that did not involve insensitive belief.[5]

I do not think that these are good reasons to reject sensitivity as a necessary condition for knowledge. I shall suggest in this chapter that, contrary to what is typically assumed, a sensitivity requirement for knowledge does *not* license abominable conjunctions, and that the alleged counterexamples in the literature are no such thing. The Lewisian contextualism described in the previous chapter will play central roles. I do concede that sensitivity doesn't 'solve the Gettier problem' in the sense

[1] Following a standard, if regrettable, philosophical convention, I use 'counterfactual conditionals' broadly to refer to subjunctive condtionals generally, whether or not the antecedent is 'counter to fact'. See Bennett (2003, pp. 11–12).

[2] Some contemporary epistemologists deny that Chloe can have a justified belief in a case like this, because they think that justification *requires* knowledge. I here assume, with epistemological orthodoxy, that such views are mistaken. I'll discuss this approach in detail in 4.3.

[3] Nozick (1981) and Dretske (1970) each offer versions of such a requirement. As will emerge in §2.6 below, there are various ways to make precise the counterfactual sensitivity condition. See also Roush (2005) for a more contemporary defence of a sensitivity condition that is not given in terms of counterfactuals.

[4] Sosa (1999) was especially influential on this score.

[5] Kripke (2011) gives one influential such case.

of figuring into an exceptionless analysis of knowledge, but a property needn't figure into an analysis in order to be a genuinely explanatory condition.[6] In my view, the deep connections between knowledge and counterfactual conditionals that were posited by Nozick and Dretske are genuine and informative ones. But they cannot be properly recognised or appreciated without contextualism.

Given the kind of contextualism I favour, I believe that a sensitivity requirement on knowledge is defensible. Furthermore, the connections that will emerge between knowledge and counterfactual conditionals will, I think, further confirm the knowledge-first idea that my contextualist approach to knowledge renders both knowledge itself and the particular contextualist parameters gestured at in Ch. 1 to be of theoretical interest.

I'll begin in this chapter by defending a particular approach to counterfactuals, motivated by a puzzle analogous to the skeptical paradox. Then I'll turn to an articulation and defence of a sensitivity condition on knowledge.[7]

2.1 Two Puzzles

Just as there is a skeptical paradox for knowledge ascriptions, so too is there a parallel puzzle about counterfactuals. Consider this instance of the skeptical paradox: Blanche has told Ida that there will be a lecture on logic today. Blanche is telling the truth, and Ida, having no particular reason to doubt her, believes her. Under the circumstances, (1) and (2) below are each intuitive, but their conjunction, (3), is under the circumstances abominable:[8]

(1) Ida knows that there will be a lecture on logic today.

(2) Ida doesn't know whether Blanche is lying.

(3) Ida doesn't know whether Blanche is lying, but she knows that there will be a lecture on logic today.

As sketched in Chapter 1, contextualism about 'knows' offers an attractive solution to this puzzle. Sentence (1) says that Ida satisfies some relatively modest epistemic standard with respect to the proposition that there will be a logic lecture; sentence (2) says she doesn't satisfy a more demanding epistemic standard with respect to the proposition that Blanche isn't lying. But in no context need we accept (3); (1) and (2) are each true in *a* context, but in no context are they both true.

Here is an analogous puzzle about counterfactuals. Gama is the king; Arac, his eldest son, is first in the line of succession. Guron is Gama's second son. So this counterfactual is very plausible:

(4) If Gama were to be killed tonight, Arac would be crowned tomorrow.

Notice, however, that there are many possibilities in which Gama is killed but Arac isn't crowned. For example, Arac might also be killed tonight, in which case Guron, not Arac, would be crowned. In particular:

(5) If Gama were to be killed tonight, and Arac were also killed tonight, Arac would not be crowned tomorrow.

But the conjunction of the ordinary (4) with the 'skeptical' (5) is abominable:

[6] Compare Jenkins (2006) and Williamson (2000, p. 32).

[7] The central ideas of this chapter are a redevelopment, extension, and defence of ideas first put forward in Ichikawa (2011b).

[8] The apt 'abominable' for conjunctions such as these is from DeRose (1995, pp. 27–8).

(6) If Gama were to be killed tonight, and Arac were also killed tonight, Arac would not be crowned tomorrow, but if Gama were to be killed tonight, Arac would be crowned tomorrow.

So as in the case of the skeptical paradox about knowledge, we have a puzzle. We have two claims that are each intuitively plausible, but whose conjunction is abominable. How should we proceed?

DeRose (1999) discusses a similar puzzle, also relating it to the skeptical paradox. Like my puzzle, DeRose's trades on the tension between ordinary counterfactuals like (4) and the recognition of the possibility of cases where the antecedent holds and the consequent fails. Unlike mine, DeRose's version codifies that recognition as a might-counterfactual. Adapting his case to fit my example, DeRose's puzzle is:

(7) If Gama were to be killed tonight, Arac would be crowned tomorrow.

(8) If Gama were to be killed tonight, Arac might not be crowned tomorrow.

(9) If Gama were to be killed tonight, Arac might not be crowned tomorrow, but if Gama were to be killed tonight, Arac would be crowned tomorrow.

As in (4)–(6) above, the problem is that the first two claims each seem intuitively true, but their conjunction is abominable. DeRose's solution to his version of the puzzle involves interpreting the 'might' in (8) and (9) as an epistemic operator, rather than as the dual of the metaphysical 'would' in (7). In other words, DeRose suggests that the relevant reading of (8) is one that says that *it is consistent with what we know* that if Gama were killed tonight, Arac would not be crowned tomorrow. The abominability of (9), on this approach, can be explained, not by inconsistency, but by its Moore-paradoxicality. It is a familiar point that in general, sentences of the form 'p but it might be that not-p' or 'p but I don't know that p' are infelicitous, even though possibly true. The standard explanation for this phenomenon is that the admission that one doesn't know that p is tantamount to allowing that one is not qualified to assert that p.[9]

DeRose's solution seems to me to do a fine job with the (7)–(9) puzzle he discusses. But it does not generalise to the (4)–(6) puzzle with which I began, for the straightforward reason that none of (4)–(6) contain an instance of 'might'. So interpreting 'might'-counterfactuals as epistemic is irrelevant for (4)–(6).

Note that one could also construct a parallel version of the skeptical paradox using 'might' language.

(10) Ida knows that there will be a lecture on logic today.

(11) Blanche might be lying.

(12) Blanche might be lying, but Ida knows that there will be a lecture on logic today.

(Recall that Blanche's testimony is Ida's only source of evidence.) As in the case of (1)–(3), the first two premises each sound intuitive, but their apparent conjunction is abominable. But one needn't embrace contextualism to resolve this version of the puzzle; a move parallel to DeRose's move with respect to (7)–(9) is available. We most naturally read (11) as indicating that it is consistent with what *we* know that Blanche is lying; but this is straightforwardly consistent with

[9] Moore (1962/2013, p. 277). See Ch. 6.

(10), which is a claim about *Ida*'s epistemic position, not ours. So (12) is consistent, but Moore-paradoxical; the first conjunct implies that the speaker isn't in a position to assert the second conjunct. But no similar move can be used to resolve the (1)–(3) paradox, which involves no instance of 'might'. The infelicity of (3), like that of (6), cannot be explained by Moore-paradoxicality.

So in both cases, I choose to focus on the form of the puzzle that does not involve 'might' operators. I embrace contextualism to solve the (1)–(3) puzzle; what about (4)–(6)? The options are parallel: we can deny a premise, accept the conjunction, or embrace contextualism. Let us consider the options in turn.

First, the choice corresponding to the skeptical resolution of (1)–(3) is to deny the truth of (4). Maybe after all it's false that if Gama were killed tonight, Arac would be crowned tomorrow. Reflection on the kinds of possibilities in which we're trading can certainly motivate one to go this way. The cost to intuitive verdicts, however, is high; it is straightforward to extend the puzzle to nearly all ordinary counterfactual sentences, many of which are intuitively true. Alan Hájek's monograph in progress, *Most Counterfactuals are False*, takes just this line, for roughly this reason.[10] Like radical skepticism, this is an extreme response to the puzzle.

Second, we could deny (5). This would be pretty hard to stomach, as a solution to the problem. Premise (5) says that if Gama and Arac were both killed tonight, Arac wouldn't be crowned tomorrow. But it is not at all plausible under the circumstances that Arac would be crowned tomorrow after being killed tonight. So denying (5) does not look like a promising strategy.[11] I find it notable that the absurd denial of (5) corresponds to the widespread Moorean response to the skeptical paradox—denying (2). Insofar as one is impressed with the parallel nature of these puzzles, one might be tempted to judge the Moorean response less satisfying for the company it keeps.

Third, one might embrace the abominable conjunction (6). This also represents a high intuitive cost; but it is one we need not pay, even if we have accepted (4) and (5). Instead, we can embrace a contextualist reading of the relevant counterfactuals. In §§2.2–**??**, I will spell out what seems to me a promising framework for such a reading; in §§2.6–2.9, I will argue that the approach defended, combined with my approach to knowledge given in Chapter 1, has interesting implications for traditional ideas about the relationship between knowledge and sensitivity.

I begin by considering, as a point of comparison, David Lewis's famous treatment of counterfactuals.

2.2 David Lewis on Counterfactuals

According to David Lewis's influential (1973) account, counterfactuals are *variably strict conditionals*. A counterfactual is true just in case some material conditional holds in all of the relevant set of worlds. This is an appealing shape for a view of counterfactuals; my own view will develop this observation in a contextualist way: which worlds are relevant depends in part on the conversational context of the speaker. It isn't always recognised, but Lewis is a contextualist about counterfactuals too. Still, he begins by developing his variably strict conditional account in a different way. According to Lewis, to evaluate a counterfactual *if A were the case, then C would be the case*, we examine

[10] See also Hawthorne (2004, p. 5, fn. 10).

[11] What exactly would follow from the denial of (5) depends on controversial matters about how counterfactuals interact with negation. Defenders of 'conditional excluded middle' like Stalnaker (1981a) and Williams (2010) will hold that denying (5) amounts to affirming that if Gama and Arac were killed, Arac would be crowned. Others, like Lewis (1973) and Kratzer (2012) will think the denial weaker than that opposing counterfactual; still, it is strong enough to be implausible. For example, in Lewis's framework, denying (5) amounts to admitting that among the nearest worlds where both Arac and Gama are killed tonight, there is at least one where Arac both dies tonight and is crowned tomorrow. So one should reject (5) whether or not one accepts conditional excluded middle. Only someone like Hájek, who thinks on different grounds that almost all counterfactuals are false (and so has a different solution to the puzzle) should be interested in denying (5).

the set of A-worlds that differ less than any non-A-worlds from the actual world, and check whether C is true in all of those worlds.

This is meant to explain some apparent logical features of counterfactuals. For instance, it explains why counterfactuals are not closed under strengthening of the antecedent, the way that the material conditional is:

> $A \supset C$ entails $(A \,\&\, B) \supset C$

But not

> $A \mathrel{\square\!\!\rightarrow} C$ entails $(A \,\&\, B) \mathrel{\square\!\!\rightarrow} C$

My (4) and (5) provide an apparent counterexample to the latter: if Gama died, Arac would be crowned; but not: if Gama and Arac died, Arac would be crowned. Lewis's explanation is that, although both counterfactuals are strict conditionals, they are strict conditionals over different domains. The domain for (4) includes all the nearest worlds where Gama dies; the domain for (5) includes the nearest worlds where both die. On the assumption that all worlds where both die are further from the nearest worlds where only Gama dies, these are distinct—indeed, non-overlapping—domains. So Lewis comfortably handles the intuitive truth of (4) and (5). But Lewis does not offer any obvious means to resist (6), their abominable conjunction.

Kit Fine (1975) famously objected to Lewis's invocation of similarity in this context. Fine pointed to cases in which intuitively, counter-to-fact possibilities would have had a dramatic effect on the course of history; but in such cases, there are possible worlds where, due to the presence of additional factors, the difference from actuality would be minimised. For example,

> [t]he counterfactual 'if Nixon had pressed the button there would have been a nuclear holocaust' is true or can be imagined to be so. Now suppose that there never will be a nuclear holocaust. Then that counterfactual is, on Lewis's analysis, very likely false. For given any world in which antecedent and consequent are both true it will be easy to imagine a closer world in which the antecedent is true and the consequent false. For we need only imagine a change that prevents the holocaust but that does not require such a great divergence from reality. (Fine, 1975, p. 452)

In response, Lewis (1979a) clarified that his invocation of 'similarity' was not meant to be understood in this simple way; certain kinds of divergence from actuality—'miracles' and 'quasi-miracles'—are much more significant than others, such as whether there is a nuclear holocaust. The worlds Fine imagines, where Nixon presses the button, but the button malfunctions, are in Lewis's particular intended sense 'less similar' to actuality than worlds where Nixon presses the button, resulting in nuclear holocaust. As a preface to this project, however, Lewis admits that he is attempting to articulate only one notion of similarity that we make use of in evaluating counterfactuals; some counterfactuals call for different similarity relations.

Of particular interest for my project is the suggestion that among the factors that influence what kind of similarity relation is relevant are facts about the conversational context. In what follows, Lewis means 'context-sensitivity' by 'vagueness'. He writes:

> What is going on, I suggest, can best be explained as follows. (1) Counterfactuals are infected with vagueness, as everyone agrees. Different ways of (partly) resolving the vagueness are appropriate in different contexts. Remember the case of Caesar in Korea: had he been in command, would he have used the atom bomb? Or would he have used catapults? It is right to say either, though not to say both together. Each is true under

> a resolution of vagueness appropriate to some contexts. (2) We ordinarily resolve the vagueness of counterfactuals in such a way that counterfactual dependence is asymmetric. ... (3) Some special contexts favor a different resolution of vagueness, one under which the past depends counterfactually on the present and some back-tracking arguments are correct. If someone propounds a back-tracking argument, for instance, his cooperative partners in conversation will switch to a resolution that gives him a chance to be right. ... But when the need for a special resolution of vagueness comes to an end, the standard resolution returns. (4) A counterfactual saying that the past would be different if the present were somehow different may come out true under the special resolution of its vagueness, but false under the standard resolution. If so, call it a *back-tracking counterfactual*. Taken out of context, it will not be clearly true or clearly false. Although we tend to favor the standard resolution, we also charitably tend to favor a resolution which gives the sentence under consideration a chance of truth. (Lewis, 1979a, p. 34)

Lewis is endorsing a form of contextualism about counterfactuals. Which notion of similarity is relevant may change according to the conversational context; therefore, whether a counterfactual sentence is true depends not only on the antecedent, the consequent, and the state of the world, but also upon the conversational context. When faced with a counterfactual, we imagine a scenario described by the antecedent. But just what we are to imagine is underdetermined by the antecedent alone; *how* are we to imagine Caesar in command in Korea? Should we imagine him having replaced General MacArthur? Or should we imagine the Roman Empire having expanded to East Asia?

Once we are convinced that this sort of limited contextualism is true about counterfactuals, it is worth investigating whether such context-sensitivity runs much deeper—and what sorts of general principles might underwrite all counterfactuals.

2.3 Counterfactual Contextualism

Let's return to our puzzle about counterfactuals:

> (4) If Gama were to be killed tonight, Arac would be crowned tomorrow.
>
> (5) If Gama were to be killed tonight, and Arac were also killed tonight, Arac would not be crowned tomorrow.
>
> (6) If Gama were to be killed tonight, and Arac were also killed tonight, Arac would not be crowned tomorrow, but if Gama were to be killed tonight, Arac would be crowned tomorrow.

The knowledge contextualist resolves the knowledge puzzle by suggesting that 'know' invokes a different epistemic standard in each premise; the counterfactual contextualist, similarly, admits that (4) and (5) each express a truth, but that changing features of the conversational context prevent the apparent entailment from (4) (in its context) and (5) (in its context) to (6) (in its context).

One bit of support for this approach is that pairs of counterfactuals like these appear to be sensitive to dynamic conversational features.[12] It sounds bad to offer (4) and (5) in the same breath; but the order in which they're presented makes a difference as to *how* bad it sounds. In particular, (7) sounds much better than (8):

[12]There are noncontextualist strategies for this kind of data too—of particular note is the pragmatic treatment offered in Moss (2012), according to which Heim sequences are merely pragmatically infelicitous, rather than contradictory. I agree with Lewis (2015) that a semantic treatment is preferable.

> (7) If Gama were killed tonight, Arac would be crowned tomorrow. But if Gama and Arac were killed tonight, Guron would be crowned tomorrow.
>
> (8) If Gama and Arac were killed tonight, Guron would be crowned tomorrow. But if Gama were killed tonight, Arac would be crowned tomorrow.

It's debatable whether 'Heim sequences' like (7) are consistent[13]—perhaps they are, or perhaps the second conjunct amounts to an implicit retraction of the first. But at least we can make sense of someone uttering (7). In contrast, it is difficult to see any context in which (8) is comprehensible; it seems flatly contradictory.

Something like the same pattern emerges for knowledge ascriptions; (9) is preferable to (10):

> (9) Ida knows that there will be a lecture on logic today, but she doesn't know whether Blanche is lying.
>
> (10) Ida doesn't know whether Blanche is lying, but she knows that there will be a lecture on logic today.

Remember, Ida's only source of evidence here is Blanche's testimony. So (9) and (10) each amount to an abominable conjunction. Still, (9) is closer to acceptable than (10) is; as it has often been observed, it is easier to raise standards than to lower them. Given the Lewisian framework developed in Chapter 1, a knowledge ascription of the form 'S knows that *p*' is true only if S's evidence eliminates 'all' not-*p* cases, where different cases count as among 'all' the cases in different conversational contexts. The first conjunct of (9) is only true under the circumstances if we are ignoring the possibility that Blanche is lying; this is difficult to do after explicitly mentioning it, which is why (10) is so bad.

Inspired by this parallel, I propose this way of articulating truth conditions for counterfactuals:

> A counterfactual of the form 'If A were the case, C would be the case' is true just in case 'all' of the possibilities in which A is true are possibilities in which C is true.

This 'all', like the one in Lewis's account of knowledge, takes a context-sensitive scope. So, we might say, following Lewis:

> A counterfactual of the form 'If A were the case, C would be the case' is true just in case all of the A possibilities are C possibilities (Psst!—except those possibilities we're properly ignoring).

It is interesting that Lewis briefly considers a contextualist view of this kind in *Counterfactuals*, dismissing it as 'defeatist,' writing that it 'consigns to the wastebasket of contextually resolved vagueness something much more amenable to systematic analysis than most of the rest of that mess in that wastebasket.' Lewis (1973, p. 13) As I argued in the Introduction to this book, it is flatly fallacious to suppose that, in offering a contextualist semantics for discourse about a given topic, one thereby commits oneself to the idea that the topic in question is uninteresting or less fundamental. So it is not at all clear why Lewis thinks such 'wastebasket' terminology is appropriate. I do not. While I don't share Lewis's commitment to the idea that counterfactuals play central roles in metaphysics, I also don't think that the contextualist semantics offered here is any obstacle to it. After all, I also offer a contextualist semantics for 'knows', and I *am* committed to the idea that knowledge plays central metaphysical roles.

[13] I follow Karen Lewis (2016) in calling these 'Heim sequences'; they entered the literature in von Fintel (2001, p. 130), which cites an MIT presentation by Irene Heim in 1994.

Lewis's own reservations notwithstanding, a contextually variable strict conditionals analysis of counterfactuals is every bit as plausible as a contextually variable infallibilist analysis of knowledge. It is such a view that I wish to develop. As will emerge, the result is interesting, not only because of the parallel issues between knowledge and counterfactuals, but also because it vindicates more direct connections that have sometimes been posited between the two.

2.4 Rules for Possibilities

As I discussed in §1.7, Lewis's attempt to spell out the rules—to determine a function from contexts and indexes to sets of possibilities relevant for a given knowledge ascription—was overly ambitious. Nevertheless, some of his rules did gesture helpfully at features of context and index that play roles in settling the truth of a knowledge ascription; by understanding them, not as a theory with predictive power, but as a gesture at some of the considerations at play, they can help us better to understand how the dynamic features of discourse behave. For example, Lewis's 'Rule of Actuality', which says that the world of the index is always relevant, ensures the factivity of knowledge ascriptions in all contexts; the 'Rule of Attention', which says that any possiblity considered in the context is relevant, while far from literally true, at least gestures at the plausible idea that the speaker's attention and communicative intentions can make a difference for what it takes to satisfy 'knows' in a context.

In a similar way, I think it's instructive to gesture at broad principles that will influence which possibilities are relevant for the purpose of a counterfactual. As in the case of those possibilities relevant for knowledge ascriptions, I have no ambition to produce a precise set of rules, but some constraints and gestures will help clarify the kind of contextualism I am proposing.

I say that a counterfactual A $\boxrightarrow$ C is true just in case all relevant A possibilities are C possibilities. Which possibilities are relevant depends in part on the context. Notice, however, that the view will not be particularly plausible if the relevant possibilities depend *only* on the context. Otherwise, for reasons very similar to those discussed in §1.10, counterfactuals couldn't be sensitive to features of the index of which the speakers are unaware. For example, suppose that in context c_1, which takes place in world w_1, Hildebrand says:

(11) If Gama were to arrive without the Princess, I'd put him in prison.

Hildebrand is speaking truly. In fact, as it turns out (though he doesn't know it yet), in w_1, Gama *will* arrive without the Princess, and Hildebrand *will* put him in prison. But consider w_2, where things are different. In w_2, Gama will arrive with a tremendous invasion force, but again without the Princess. In this world, Gama sacks Castle Hildebrand, and it is Hildebrand, not Gama, who is put in prison. So in w_2, Gama arrives without the Princess, but Hildebrand does not put Gama in prison. It seems clear to me that this suffices for the falsity (in w_2) of what Gama said (in c_1 in w_1). What he said was true where he said it—w_1—but it's false in a world where Gama's Princessless invasion sacks the castle. Since none of the facts about whether a given world is a world where Gama arrives without the Princess, or Hildebrand puts Gama in prison, vary in truth relative to different worlds, the only way to make sense of this is to suppose that index, as well as context, plays a role in determining which worlds are relevant. In this case, w_2 is relevant when (11) is evaluated relative to c_1 and w_2, but it is irrelevant when (11) is evaluated relative to c_1 and w_1.[14]

So the index can make a difference for which worlds are relevant. In particular, the lesson of this case appears to be that the world of the index itself is always relevant. The counterfactual says that

[14] I am assuming that c_1 is a relatively 'low-standards' context, where possibilities like w_2 are properly ignored. On the contextualist approach under consideration, there will also be more 'skeptical' possible contexts where, even evaluated relative to w_1, (11) would be false.

every antecedent-world is a consequent-world; the index is always among the worlds in question. So no counterfactual can be true relative to a world where the antecedent is true and the consequent is false. This cousin of Lewis's rule of actuality for 'knows' ensures the validity of counterfactual modus ponens.[15]

We should also say at least something about *how* conversational context affects which worlds are relevant. Consider again the puzzle I began the chapter with:

> (4) If Gama were to be killed tonight, Arac would be crowned tomorrow.
>
> (5) If Gama were to be killed tonight, and Arac were also killed tonight, Arac would not be crowned tomorrow.

On the contextualist resolution to the puzzle offered, counterfactuals like (4) are context-sensitive: in more lax contexts in which the possibility that Arac is killed tonight is properly ignored, (4) expresses a weaker claim than it does in contexts where that possibility is considered live. So something like this principle is suggested: possibilities attended to in the context are relevant. Whether (4) expresses a truth in the world in question depends in part on the speakers' attitudes towards the possibility that Arac will be killed tonight. If they are ignoring it, it may be irrelevant; if they are considering it, then it is relevant, and so (4) ends up saying something implausibly strong. There are admittedly sensitive questions about exactly what sort of attitude constitutes 'attention' in the relevant sense; but this is exactly as it is in the case of relevant possibilities for 'knows'.

In fact, the connection between the possibilities relevant for counterfactuals and those relevant for knowledge ascriptions is deeper than mere parallel; at least with regard to what I've considered so far, it seems as if it may be *identity*. Lewis's rule of actuality, which guarantees the factivity of knowledge, *just is* the rule that the world of the index is always relevant, which is what is needed to ensure the validity of counterfactual modus ponens. And the much more nebulous rule of attention, which gives Lewis's 'knows' its contextualist character, and accounts for resolving the skeptical paradox, looks to have exactly the same profile of the role that attention plays in the (4)–(5) puzzle. At this stage, it is beginning to look like a plausible working hypothesis that it is the *very same domains* that are relevant for knowledge ascriptions and counterfactuals alike. The hypothesis is confirmed by consideration of particular cases, where consideration of ignorable possibilities that would falsify counterfactuals and knowledge ascriptions alike.

For example, Blanche is, and has long been, Professor of Abstract Science. This, as Melissa is well aware, won't change any time soon. So (12) is true:

> (12) Melissa knows that Blanche will be Professor of Abstract Science tomorrow.

In the unlikely event that Ida, the principal, abandoned her position tonight, Blanche would leave her post as Professor of Abstract Science and take her place:

> (13) If Ida resigned tonight, Blanche would be principal tomorrow.

The domain for (12) and (13), in their natural, 'nonskeptical' contexts, does not include cases in which Blanche resigns tonight altogether. (These cases are very distant and legitimately ignored; Blanche is not likely to leave the university under any circumstances.) But she *could* resign tonight.[16] We may explicitly bring those cases into the domain; doing so undermines both sentences:

[15] Proof: Suppose there is a counterexample to counterfactual modus ponens. That is, suppose there is a world w_x where A is true, and A $\Box\!\!\rightarrow$ C is true, and C is not true. Since A $\Box\!\!\rightarrow$ C is true, every relevant A world is a C world. Since w_x is an A world that is not a C world, it must not be relevant; but this is just what the constraint in question precludes.

[16] The modal 'could' is also context-sensitive; this is plausibly false in contexts in which (12) and (13) are true. See Ichikawa (2016b).

(14) Blanche could resign tonight, but Melissa knows that Blanche will be Professor of Abstract Science tomorrow.

(15) Blanche could resign tonight, but if Ida resigned tonight, Blanche would be principal tomorrow.

Both of these sentences are false in their contexts—each is undermined in just the same way, by the consideration of a particular counterexample: a case where Blanche resigns tonight. It is a counterexample to the knowledge ascription because it's a case where Melissa's evidence is as it actually is, but Blanche will not be Professor of Abstract Science tomorrow; it is a counterexample to the counterfactual utterance because it's a case where Ida resigns tonight, but Blanche is not principal tomorrow.

So I think there is a decent prima facie case to be made, based on reflection on cases, that the domains of relevance for counterfactuals may be the same as those for knowledge ascriptions. Another, broader motivation for positing such an identity is the theoretical simplicity of the resulting picture. Indeed, one might well think that the fact that these kinds of contextually-determined domains have *multiple* metasemantic roles to play in settling context-sensitive discourse provides greater motivation to posit them in one case or the other. Suggesting there is a contextual parameter that plays a central role in giving content to knowledge ascriptions, but that this is its *only* role, may feel ad hoc; so likewise for the corresponding claim for counterfactuals. But discovery that posits are well-suited to play multiple theoretical roles is an excellent reason to believe that they're on to something. A third positive reason to treat the relevant domains as coextensive is that, as I'll argue in §2.6, doing so permits attractive explanations for the apparent necessary conditions of safety and sensitivity on knowledge.

First, however, I'll consider an objection to my proposed relationship between knowledge's domains and counterfactuals'.

2.5 Karen Lewis and Ignorance of Counterfactuals

Here is an objection due to Karen Lewis.[17] According to my proposal, a counterfactual conditional requires that 'all' worlds in which the antecedent is true are worlds in which the consequent is true, where the 'all' in question also determines the relevant alternatives for a given knowledge ascription in the same context. This suggests a prima facie surprising connection between metaphysical and epistemic modality. Here, loosely speaking, is a source of worry: the truth of counterfactuals is a *metaphysical* matter, while the truth of knowledge ascriptions is an *epistemic* matter. Connecting their truth conditions in this way threatens either to make counterfactuals rely surprisingly on epistemic considerations, or to make knowledge ascriptions depend in a surprising way on metaphysical facts.

One can also precisify this vague worry into a potential counterexample. Consider an ornamental box. From its visual appearance, one can't really tell what it's made of or how hardy it is; but as a matter of fact, it's quite fragile. In particular, if someone pushed it off the table, it would break. (I am presently writing in a context in which strange possibilities, like those possibilities in which the dog happens to drag a pillow under the table at just the right moment, are ignored.) Nevertheless, it seems clear that someone—Florian, for instance—might be ignorant of this true counterfactual. Florian might have no evidence, or even misleading evidence, about the durability of the box. Moreover, an ascription of such ignorance looks like it might be true even in those non-skeptical contexts in which the counterfactual is true: for some possible boxes, background conditions, and epistemic states, the following might both be true in a given context:

[17] It appears in a footnote in Lewis (2016), and in more detail in Lewis (2017).

(13) If someone pushed the box off the table, the box would break.

(14) Florian doesn't know that if someone pushed the box off the table, the box would break.

But, the worry continues, it is difficult to see how on my view (14) is consistent with (13). I say the set of relevant possibilities for knowledge ascriptions and counterfactuals is the same. According to my approach to (13), it is true just in case all relevant worlds where the box is pushed are worlds where the box breaks; there are no relevant worlds where the box is pushed but remains unbroken. But if there are no such relevant worlds, then trivially, Florian's evidence cannot be consistent with any such relevant worlds. So it seems Florian *would* know the conditional in question, since his evidence eliminates 'all' worlds where it is false.

This is a strong objection. Unless one is willing to deny that (13) and (14) can be true in the same context, it (modulo fn. 18) refutes the conjunction of my suggestion here with the view of 'Elusive Knowledge', according to which satisfaction of 'knows' just *is* a matter of the evidence eliminating counterpossibilities. It also, I think, successfully refutes the letter of the view given in Ichikawa (2011a). My present subtler approach, however, given in §1.7, has more resources to avoid the objection. Lewis is right that, relative to a context in which (13) is true, there are no relevant possibilities in which someone pushes the box and it doesn't break; so there are no such uneliminated possibilities. It seems then that Florian's evidence, like *any* body of evidence, trivially establishes the conditional of which Florian appears ignorant.[18] But it doesn't follow that Florian knows the conditional; my §1.7 has it that to satisfy 'knows', Florian must properly base a belief on sufficient evidence. But under the circumstances, even though the evidence is sufficient, Florian lacks the competences it would require to base such a belief in such a way. To see how my view can accommodate such a position, consider an analogy.

Suppose that Florian's evidence conclusively establishes that there are six caps in front of him.[19] Then Florian's evidence rules out all possibilities in which Florian is faced with a number of caps that is not the only even perfect number that is not the sum of successive odd cubes.[20] Nevertheless, this is a fact that Florian doesn't know, because he's unable to appreciate the connection between his evidence and this hypothesis. Knowledge doesn't *only* require sufficient evidence; it requires the properly skilful processing of such evidence. In this case, his evidence *does* entail the hypothesis in question, but he lacks the ability to capitalise on this fact. Another case of this form would

[18] Here is an additional complication: that the evidence entails the conditional doesn't *directly* follow from the setup of the case. Given the semantics I've been working with, any time all relevant A-worlds are B-worlds, the conditional A $\Box\!\!\rightarrow$ B is true. It *isn't* trivial, however, that any time all relevant A-worlds are B-worlds, *all relevant worlds* are worlds in which A $\Box\!\!\rightarrow$ B is true. As I argued in §1.8, context and index each have a role to play in determining which possibilities are relevant. So, even if, relative to the set of relevant possibilities established by the actual world and context C, all A-worlds are B worlds, there might be some relevant world *w* such that, among to the set of relevant alternatives established by C *and w*, there are some (A & ~B) worlds.

So in the case we've been considering, it may be that there are no relevant worlds where the box is pushed without breaking, but there *are* some relevant worlds where the conditional is false, because relative to *those* worlds (and the actual context) there are worlds where it's pushed without breaking.

I don't think this provides a deep solution to Lewis's worry, however. The worry goes beyond the charge that Florian might not know the counterfactual; suppose for instance that the counterfactual is true, and moreover, that its antecedent is true: the box has been pushed off the table, and that Florian knows that. It seems that Florian may still not be in a position to know that the box will break, even though all relevant worlds compatible with his evidence are ones in which it will break.

[19] Given David Lewis's phenomenal conception of evidence, this will only be plausible relative to certain not-too-skeptical conversational contexts. We may assume that we're working in one such. (Or we may work with a richer conception of evidence; see Ch. 3.)

[20] A perfect number is a number that is equal to the sum of its proper positive divisors. Six is a perfect number because 1 + 2 + 3 = 6. All other even perfect numbers are sums of successive odd cubes. For example, $28 = 1 + 2 + 4 + 7 + 14 = 1^3 + 3^3$.

be a version of the 'Problem of the Speckled Hen'—suppose that Florian's experience represents twenty-two caps, but he hasn't counted them.[21] This would also be a case where Florian's evidence conclusively establishes something—that there are twenty-two caps—but where Florian lacks the cognitive resources to know it.

I wish to assimilate Florian's ignorance about the box's fragility to cases of this sort. Relative to the conversational contexts in which the counterfactual (13) is true, Florian *does* have evidence that establishes the conclusive relationship between the box's falling and its breaking; but he isn't in a position to appreciate the connection. Perhaps some readers will be surprised by the fact that, on my view, such 'hidden' features of the world as the box's composition can influence which possibilities matter for the purpose of a knowledge ascription (and so matter for whether the evidence is relevantly conclusive); but on reflection, I think this is the kind of conclusion that everyone ought to accept. For example, all non-skeptics who endorse orthodoxy about Gettier cases must allow that whether e.g. the possibility that this is a fake barn is relevant (and so whether the evidence is conclusive for its being a barn) can depend on the subject's hidden environment.

And although I agree with Lewis that there are *some* contexts where (13) and (14) are both true, there are also contexts relative to which (13) is false because it receives a more epistemological reading—that is to say, there are contexts in which there are relevant possibilities in which the box is *not* fragile. Consider for instance a *backtracking* context—suppose it is known that Ida, who is firmly in control of the box, would only allow it to be pushed if she knew that it was robust enough to survive the fall.[22] In a context in which this background is taken for granted, an utterance of (C′) might express a truth:

> (15) If someone pushed the box off the table, the box would be tough enough to survive the fall.

In such a context (13) would be false.[23] So I do think that some contexts generate more epistemic readings of counterfactuals like (13).

I conclude that there is no definitive objection to a suitably nuanced Lewisian contextualism that identifies knowledge-relevance with counterfactual-relevance. An additional advantage to this theoretically attractive picture is that it allows vindication of attractive relationships between knowledge, sensitivity, and safety. I turn now to those connections.

2.6 Knowledge and Sensitivity

Counterfactual conditionals seem intuitively to have a connection to knowledge. The twentieth century saw attempts to codify this intuitive connection as a necessary condition on knowledge. Here is the condition I am interested in:

> Sensitivity: S's belief that p, on the basis of evidence e, is *sensitive* if and only if this counterfactual is true: if p were not the case, S would not believe that p on the basis of e.

[21] See Ch. 1, fn. 31.

[22] On backtracking counterfactuals generally, see Lewis (1979a, p. 456). For the emphasis on epistemic readings of counterfactuals, see Edgington (2011, pp. 238–9).

[23] So although I agree with much of the spirit of Lewis's (Lewis, 2015, p. 8) remarks about the objectivity of possibilities' relevance—I agree, for instance, that it is not *solely* a matter of conservational intents and interests—I think she goes too far in her Lewis (2016, p. 29) suggestion that in cases like this one, the possibility that the box falls and doesn't break can *never* be irrelevant. (Whether my stance is consistent with the letter of Lewis (2015, p. 8) depends on how exactly one interprets the latter's discussion of 'high-probability' events.)

Here is an interesting and influential claim:

> SK: Sensitivity is a necessary condition for knowledge.

SK has at least two virtues. First, in many cases that approximate knowledge but fall short, the beliefs in question are insensitive. So SK seems to explain some important data. Second, SK itself has intuitive appeal. Setting aside judgments about particular cases, there is something cogent about the general skeptical argument form: you'd think that p even if p were false, so you can't know that p. Given these virtues, SK does seem to have at least something going for it.

Few epistemologists today embrace SK. (And among the few exceptions, some, such as Roush (2005), make use of alternate formulations of sensitivity that eschew counterfactual conditionals.) I think the main reason that SK has fallen out of favour is that it is thought to be inconsistent with certain obvious cases where we should easily be able to know deductive consequences of other things that we know.[24] For example, it seems that no one who knows that she is looking at a red wall can be ignorant about whether she's looking at a white wall illuminated with red light. This is precisely the same pattern of intuitions that motivates the skeptical argument from ignorance discussed in §2.1 above. Schematically, the problem is this: for many *o*rdinary propositions o that are intuitively known, and for which a subject has a sensitive true belief, there will be *s*ceptical propositions s such that: (1) the subject does not have a sensitive belief that $\sim s$, and (2) $\sim s$ is an obvious deductive consequence of o. If so, endorsing the nonskeptical intuition concerning o along with SK commits one to abominable conjunctions of the form *the subject knows that o, but doesn't know that not-s*.[25]

A second, related reason that sensitivity is no longer widely regarded as necessary for knowledge is that there are intuitive counterexamples: cases of true beliefs that intuitively do amount to knowledge, but are insensitive. Notice that this is not the case with the class of objections in the previous paragraph; the particular insensitive beliefs cited *are* intuitive cases of ignorance—SK gets things right that far—but it's argued that accepting the intuition is problematic. But there are also cases where SK predicts ignorance, where this is arguably the *wrong* intuitive result. Here is an influential one from Ernest Sosa:

> On my way to the elevator I release a trash bag down the chute from my high rise condo. Presumably I know my bag will soon be in the basement. But what if, having been released, it still (incredibly) were not to arrive there? That presumably would be because it had been snagged somehow in the chute on the way down (an incredibly rare occurrence), or some such happenstance. But none such could affect my predictive belief as I release it, so I would still predict that the bag would soon arrive in the basement. My belief seems not to be sensitive, therefore, but constitutes knowledge anyhow, and can correctly be said to do so. Sosa (1999, pp. 145–46)

[24] Many epistemologists have defended much stronger principles than this—for example, a defender of *single-premise closure* will hold that *whenever* one knows something, one is in a position to know that which follows deductively from it. (A defender of *multi-premise closure* goes further still in the obvious way.) Any such principle—even the general statement of single-premise closure—is controversial. (See e.g. Lasonen-Aarnio (2008) for a summary of challenges to multi-premise closure and an extension of them to single-premise closure; see Heylen (2016) for an argument that plausible formulations of closure should relate *knowledge* to being *in a position* to know.) But note that the problem for sensitivity theorists is stronger than incompatibility with such general principles; it is that it predicts certain *particular* counterexamples which themselves are intuitively the wrong result. Note for example that the problem of abominable conjunctions has nothing to do with the aggregation of risk of error; therefore, one might deny closure *generally*, but still find the apparent consequences of SK unacceptable. So I do not think that, for example, the invocation of the closure principle in pressing skeptical arguments in Pritchard (2002, p. 101) is the best way of drawing out the tension. See §3.5, fn. 13.

[25] DeRose (1995, pp. 27–9) gives a canonical statement of this objection.

As mentioned in the introduction to this chapter, a third reason SK has fallen out of favour concerns its failure to 'solve the Gettier problem', as there are apparent Gettier cases that involve sensitive belief.[26] This not an objection to the *truth* of SK; since SK, if true, would be of substantive epistemological interest, even without figuring into an 'analysis of knowledge', I set this consideration aside. The two main objections I wish to consider in this section, then, are (1) that SK commits a non-sceptic to abominable conjunctions, and (2) that there are counterexamples to SK, like the garbage chute case. In §2.9 I will give a positive argument in favour of SK.

The objections to SK given above assume invariantist approaches to both knowledge ascriptions and counterfactuals. How do they fare when shifted into a framework like the one I am advocating in this chapter, according to which each is context-sensitive? As we have effectively seen already in Chapter 1, contextualism about 'knows' allows one to finesse the kind of 'argument from ignorance' that appears to commit SK to abominable conjunctions. That argument began with an ordinary knowledge ascription made in a non-skeptical context—the subject knows that she is looking at a red wall—and then moved on to consider a skeptical scenario—the subject is looking at a white wall illuminated by red lights. This is just the sort of situation in which a contextualist will suppose the epistemic standards have been raised.

Does the contextualist who goes this way reject SK? It depends on how exactly SK is understood. I presented SK above as an object-level claim about knowledge; a contextualist might interpret it in various ways. There *would* still be a commitment to the truth of an abominable conjunction, given this metalinguistic generalisation of SK:

> SKg_1: That S's belief that *p* be sensitive is a necessary condition for 'S knows *p*' to be true in any context.

Consider a low-standards context c_1 in which 'Melissa knows that she is looking at a red wall' is true (even though she hasn't inspected the lighting). Mention of the possibility of misleading lighting would shift to a more skeptical context; nevertheless, we can consider whether Melissa satisfies the content of 'Melissa knows that she's not looking at a white wall with red lights' relative to c_1. And on the orthodox, invariantist conception of the sensitivity counterfactual, SKg_1 demands that we treat this as false. For her belief to this effect is insensitive: if she were looking at a white wall with red lights, she'd still believe she was looking at a red wall. (Remember, we've stipulated that she hasn't checked the lighting; she's just reasonably assuming that nothing funny is going on.) So, given SKg_1, 'Melissa knows that she's not looking at a white wall with red lights' is false in c_1. But 'Melissa knows that she is looking at a red wall' is *true* in c_1, and so, so too is their abominable conjunction. A contextualist who wishes to endorse SKg_1 needn't necessarily despair at this result; one might take solace in the fact that such abominable conjunctions are *elusive*—they aren't true when uttered. Still, there is something unsatisfying about allowing them to be true at all.

The problem with SKg_1 is starker for Sosa's garbage chute case. Even if one is a contextualist who is happy to admit that 'Ernest doesn't know that the bag will soon be in the basement' is true in contexts where one is considering the possibility that it will get stuck, SKg_1 implies much more than this: it implies that this expression of ignorance is true in *all* contexts, since the necessary condition for its embedded knowledge ascription's being true in *any* context isn't met: it's false that, were the bag not to arrive in the basement soon, Ernest wouldn't believe it would.

But if one adopts the contextualist approach to counterfactuals I have advocated in this chapter, one can avoid both problems. The contextual shiftiness of the sensitivity counterfactual can 'sway

[26] For example, Kripke (2011) argues that a failure of sensitivity is not present in all Gettier cases; however, given the approach to knowledge and counterfactual conditionals I am endorsing, whether this is right ends up being a rather subtle matter.

together' with that of the knowledge ascription.[27] So consider this alternate generalisation of the sensitivity condition:

> SKg_2: For any context *C*, the content of the relevant sensitivity counterfactual in *C* is a necessary condition for the truth of the content of the relevant knowledge ascription in *C*.

Consider how the garbage chute case fares now. Ernest places his bag into the chute. Let c_l be an ordinary, nonskeptical context, and c_h be a more skeptical context where the possibility that the bag has gotten stuck is relevant. Consider first c_h. Here, 'Ernest knows that the bag will soon be in the basement' is false on the Lewisian contextualist approach of Chapter 1, since there is a relevant possibility—one where the bag sticks—uneliminated by Ernest's evidence, but where the bag doesn't reach the basement. The sensitivity counterfactual is also false. So there is no problem for SKg_2 here. What of c_l? Here, the knowledge ascription is true—Ernest's evidence eliminates all relevant counterpossibilities. What of the sensitivity counterfactual? It is tempting to conclude that it must be false, on the grounds that it just sounds obviously false under the circumstances that, if the bag didn't make it to the basement, Ernest wouldn't think it did. But this would be a mistake. That it sounds obviously false can be explained by the fact that it would be false in any context in which it was considered. But, for reasons very similar to the reasons for which skeptical scenarios count as 'known' to be false relative to nonskeptical contexts (see §1.1), this conditional is plausibly *true* relative to a context in which no possibility of the bag getting stuck is relevant.

The counterfactual is true just in case, of the relevant possibilities, *all* of those in which the antecedent is true are ones in which the consequent are true. And the cases that we were tempted to think falsified the counterfactual were precisely those that are irrelevant in c_l. So the sensitivity counterfactual seems to be false only in those same contexts in which the knowledge ascription is false. SKg_2 seems to be in good standing.

2.7 Equivocation and Necessary Conditions

Is this a genuine vindication of sensitivity as a necessary condition for knowledge? One might think not, on the grounds that such a condition should apply even in cases in which contexts shift. After all, runs this objection, these might both be true, albeit in different contexts:

> (16) Ernest knows that the bag will reach the basement.
>
> (17) If the bag weren't to reach the basement, Ernest would still think, on the basis of the same evidence, that it would.

According to the objection I am considering, the fact that each of these is true in the context in which it would likely appear is sufficient for rejecting sensitivity as a necessary condition for knowledge. I am not sure how substantive a concern this really is—perhaps the only dispute is about what is deserving of the name 'necessary condition'—but I have heard it given as an objection often enough that it is worth taking a moment to consider it. I do think that denying that sensitivity is necessary for knowledge by asserting the knowledge ascription and denying the sensitivity counterfactual, without regard to whether contextual parameters change midway through, would be a mistake. For parallel rejections are possible for many—perhaps even all—plausibly valid inferences.

Consider, for example, this *d*ownward-*e*ntailing *p*rinciple for *q*uantifiers:

[27] This terminology comes from Lewis (1973, p. 92) and DeRose (2009, p. 99).

(DEPQ) If all Fs are H, then all Fs that are Gs are H.

I hope that DEPQ is obvious. It is both a theorem of standard first-order logic, and intuitively extremely compelling. If all warriors are brutish, then all warriors in the castle are brutish. Conversely, if some warriors in the castle aren't brutish, then not all warriors are brutish. There is, I submit, nothing fishy about this reasoning—it is as valid as it sounds; DEPQ is in excellent standing.

But one can construct what superficially seem to be counterexamples to DEPQ, if one is allowed to let the contextual parameters roam free. Cyril is eating grapes at the luncheon; his plate is full of ripe, pert grapes. Restricting his quantifier to the grapes on the plate, Cyril, in c_1, utters:

(18) All the grapes look delicious.

Earlier in the afternoon, Cyril stepped on a few grapes, mashing them into the grimy floor. Flies have set into their sticky juices; (18) is false:

(19) All the grapes on the floor look delicious.

One might try to argue, on the basis of the truth of (18) and the falsity of (19), that DEPQ is false—here is a case where all the grapes look delicious, but not all the grapes on the floor look delicious. But the fallacy here is pretty transparent. The truth of (18) depended on stipulated features of its context c_1—the 'all' quantifier was restricted in a way so that only grapes on the plate were relevant. But the truth of (19)—indeed, its ordinary intelligibility—requires a different context where other grapes are considered. So the purported counterexample relies on covert changes in context. Principles like DEPQ, when given as here in a way that uses context-sensitive terms, should not be saddled with implausible metalinguistic generality.[28] The case of Cyril and the grapes gives a counterexample only to an e*x*tremely *l*iberal metalinguistic interpretation of DEPQ:

(DEPQ_{xl}) If 'all Fs are H' is true in any context, then 'all Fs that are Gs are H' is true in every context.

But we have no counterexample to a more *p*lausible metalinguistic generalisation like this one:

(DEPQ_{p}) If 'all Fs are H' is true in context *C*, then 'all Fs that are Gs are H' is true in *C*.

In just the same way, Sosa's garbage chute case refutes only implausibly general formulations of a sensitivity condition on knowledge. Indeed, we can make a version of the same point more directly. Consider this even more obvious *r*eflexivity *p*rinciple:

(RP) If P, then P.

Here is a superficial attempt at a counterexample. I point to Princess Ida and utter

(20) She is a princess.

I say something true. Then I turn and point to Melissa, and say another truth by uttering

(21) She is not a princess.

Observing these two truths constitutes no counterexample to the plausible version of (RP). Only by shifting contexts midway through can one create the (rather weak) illusory appearance as of one. The name for such shifts is 'equivocation'; there's a reason we're taught it as a fallacy.

[28] Compare Dowell (2012, pp. 285–8) on 'dynamic modus ponens'.

2.8 Strengthening the Antecedent

In an exactly parallel way to the defence of sensitivity as necessary for knowledge, on the contextualist approach to counterfactuals given above, strengthening the antecedent is also valid. This was implicit already, but this is an appropriate occasion to make it explicit.

As indicated above, most philosophers have denied that an *a*ntecedent-*s*trengthening principle for *c*ounterfactuals is correct:

ASC. A $\Box\!\!\rightarrow$ C entails (A & B) $\Box\!\!\rightarrow$ C

Pairs of counterfactuals like (4) and (5), discussed above, are treated as providing a counterexample:

(4) If Gama were to be killed tonight, Arac would be crowned tomorrow.

(5) If Gama were to be killed tonight, and Arac were also killed tonight, Arac would not be crowned tomorrow.

But as we have seen, on the semantics offered, (4) and (5) are each context-sensitive, and (4) is only true relative to contexts in which the possibility that Arac is killed tonight is irrelevant. And relative to such a context, (5) is (vacuously) true—so there is no counterexample to ASC. Like the purported counterexamples to SK, DEPQ, and RP above, this one trades on a shift in context that we should not allow. A counterexample to the interesting version of ASC would involve counterfactuals of the form A $\Box\!\!\rightarrow$ C and (A & B) $\Box\!\!\rightarrow$ C, where, *within a given context*, the former is true and the latter is false. And (4) and (5) present no such counterexample.

In fact, they fail to be a counterexample for a very general reason; on the semantics I've defended, there can *be* no such counterexample. Given the proposed treatment of counterfactuals, A $\Box\!\!\rightarrow$ C is true in context c if and only if all relevant-in-c A cases are C cases. And (A & B) $\Box\!\!\rightarrow$ C is true in c if and only if all relevant-in-c A cases that are also B cases are C cases. That is to say, ASC is just a special case (DEPQ).

Any counterexample to ASC would *ipso facto* be a counterexample to DEPQ. And there are no counterexaples to DEPQ; there are at best some pairs of sentences where, if one quietly shifts contexts between them, one will have what superficially looks like a counterexample to DEPQ.[29]

2.9 Sensitivity, Safety, and Knowledge

The point of the previous sections generalises to other logical principles thought to fail for counterfactuals. For example, *contraposition* for counterfactuals fails on the Lewis-Stalnaker semantics.

CC: A $\Box\!\!\rightarrow$ C if and only if C $\Box\!\!\rightarrow$ A

I shall suggest below that, much like ASC, CC should be accepted in a contextualist framework. Sosa (1999) has used the supposed failure of contraposition to motivate preferring *safety* as a necessary condition for knowledge, instead of sensitivity. Sosa characterises safety as the contrapositive of sensitivity:

[29] See Brogaard and Salerno (2008) for a similar defence of the validity of strengthening the antecedent.

> A belief is sensitive iff had it been false, S would not have held it, whereas a belief is safe iff S would not have held it without it being true. For short: S's belief B(p) is sensitive iff $\sim p \mathrel{\Box\!\!\rightarrow} \sim$B($p$), whereas S's belief is safe iff B(p) $\mathrel{\Box\!\!\rightarrow} p$. Sosa (1999, p. 146)[30]

As in the case of sensitivity, it is best to relativise the safety condition to evidence or methods for the belief acquisition; otherwise there will be counterexamples of the wrong sort.[31] So, modifying Sosa's approach minimally, let's consider the idea that safety is necessary for knowledge:

> Safety: S's belief that p, on the basis of evidence e, is *safe* if and only if this counterfactual is true: if S were to believe p on the basis of e, p would be the case.
>
> SfK: Safety is a necessary condition for knowledge.

It is worthwhile to begin with a relatively obvious point: although Sosa proposes SfK as an *alternative* to the sensitivity requirement SK, there is no obvious sense in which they *need* be competitors. As mere necessary conditions on knowledge, they do not preclude one another. (Indeed, Nozick accepted both SK and SfK.) Sosa sets them up as competitors because he thinks that (a) SK faces problematic consequences, and (b) the truth of SfK can account for the intuitive appeal of SK. Sosa is engaged in an *explaining-away* project.[32]

The somewhat subtle reason Sosa thinks SfK can explain the intuitive plausibility of SK is this: safety and sensitivity are contrapositive counterfactual conditions. Because CC fails in general, the contraspositives safety and sensitivity are non-equivalent. But because the failure of CC isn't intuitively transparent, it is easy for people to make the *mistaken assumption* that safety and sensitivity are equivalent, and confuse one state for the other. One has an intuitive response to the truth of SfK, and confuses this true principle with its false counterpart SK.

On the view I have been defending, CC is true in general; counterfactuals do contrapose in the sense that, relative to any given context and index, A $\mathrel{\Box\!\!\rightarrow}$ C and C $\mathrel{\Box\!\!\rightarrow}$ A will have the same truth value. The former says that all relevant A cases are C cases; the latter says that no relevant non-A cases are non-C cases. Which cases are relevant depends only on the context and index; so if these are held fixed, the counterfactuals are equivalent. As in the case of strengthening the antecedent, apparent counterexamples trade on equivocation. Sosa offers these as an apparent counterexample to CC:

> (22) If water now flowed from your kitchen faucet, it would *not* then be the case that water so flowed while your main valve was closed.
>
> (23) If water flowed from your kitchen faucet while the main valve was closed, water would not so flow.[33]

---

[30] This is one of several formulations of safety Sosa offers. Note that this formulation assumes that a *strong centering condition* for counterfactuals, according to which the index world is uniquely nearest, fails. On a Lewis–Stalnaker semantics that accepts strong centering, for instance, all true beliefs are safe by this metric, since if the belief is true, the index world will be sole relevant world. This is not Sosa's intent.

[31] The sort in question is those in which a subject responds properly to the evidence, but there are nearby worlds (that obviously do not obtain) in which the subject proceeds irrationally. I discuss this further in Ichikawa (2011a, p. 302); see also Nozick (1981, p. 233).

[32] On the nature and significance of explaining away in general, and this argument of Sosa's as a case study, see Ichikawa (2009a).

[33] Sosa (1999, p. 152, my numbering)

The approach to counterfactuals developed in this chapter undercuts this counterexample. Sosa's first counterfactual (22) is true only in a context where the possibility that water flows despite a closed valve is irrelevant. (If that possibility were relevant, then a 'skeptical' argument like that in §2.1 would undermine the truth of (22).) But relative to any such context, (23) is vacuously true. So an utterance of (23) tends to shift to contexts in which some such possibilities are relevant.

On my view, (22) and (23) are exactly similar to (24) and (25):

> (24) None of the cases where water leaks are cases where water leaks despite a closed valve.

> (25) None of the cases where water leaks despite a closed valve are cases where water leaks.

In their most natural contexts, (24) is true and (25) is false. But, since contraposition for quantifiers is valid, (24) and (25) must have the same truth conditions in any given context.

So CC is true in general, on my view; on their counterfactual formulations, safety and sensitivity are equivalent. So SK and SfK, the respective claims that each is necessary for knowledge, are also equivalent. Furthermore, given the contextualist approach to knowledge defended in Chapter 1, the safety/sensitivity condition is genuinely necessary for knowledge; SK and SfK are true. By way of reminder, here is my characterisation of the truth conditions for knowledge ascriptions, defended and articulated in §1.7:

> S knows that p just in case, for some evidence e, (i) S believes that p, where that belief is properly based on evidence e, and (ii) all the e cases are p cases.

I have already shown that the safety and sensitivity counterfactuals are equivalent in a given context. To show that they are necessary for knowledge, then, it will suffice to show that any insensitive belief fails to meet the condition above. This is easily done. Suppose S's belief that p, based on e, is insensitive. That is to say, it is false that, were p false, S wouldn't believe p on the basis of e. So there are some relevant possibilities where p is false, but S does believe p on the basis of e. So there are some relevant possibilities where e obtains but p is false. So S does not know p. Therefore, any insensitive belief is unknown; i.e., sensitivity is necessary for knowledge.

It is no accident that twentieth-century theorists tied knowledge closely to counterfactuals; there is good reason to believe that there are deep connections between them. Even the apparently strongest such connection taken seriously—a counterfactual formulation of sensitivity as necessary for knowledge—is defensible, given an independently-motivated contextualist approach to each.

Chapter 3

Evidence

It is not particularly controversial that there is a connection between knowledge and evidence; the idea that knowledge requires that a belief be supported by the subject's total evidence is all but platitudinous. I have encoded one such version of this requirement into my Lewisian approach to knowledge ascriptions. As I laid out in §1.7 (and in Ichikawa (2011a)), an attractive approach to knowledge ascriptions is a metalinguistic generalization of this principle:

> S knows that *p* just in case, for some evidence E, (i) S believes that *p*, where that belief is properly based on evidence E, and (ii) all the E cases are *p* cases.

I highlighted and motivated some differences between this formulation and that given by Lewis (1996) in §1.7. As I pointed out there, one may implement this kind of view with any of various approaches to evidence; in this chapter, I'll examine some of the choice points more closely. Lewis himself opts for a phenomenological conception of evidence, according to which one's evidence comprises all and only the facts about one's sensory and memory experiences. I will speculate a bit about the motivation for this approach in §3.11 below, and attempt to motivate an alternative.

It is also worthwhile to think through how this kind of contextualism bears on another, much stronger, connection between knowledge and evidence—namely, the thesis that knowledge *is* what makes a given proposition part of one's evidence. Williamson (2000) famously argues for the identity of evidence with knowledge—'E=K'. This chapter has three aims. First, I will argue that E=K benefits from contextualism—there is some pressure, if one adopts E=K, to adopt contextualism about both 'knowledge' and 'evidence' (and to endorse a metalinguistically general statement of 'E=K').[1] Second, I will argue that the particular Lewisian approach to knowledge I mention above, combined with E=K, yields an attractive and plausible approach to knowledge and knowledge ascriptions. (Part of this project will be to demonstrate that the superficial appearance of circularity in that combination is in fact non-vicious.)[2] Third, I will observe that the resultant view highlights an important choice point regarding basic evidence, and articulate a particular form of 'Moorean contextualism' that has so far gone unnoticed and unexplored in the literature.

3.1 Motivation for E=K

Williamson (2000, Ch. 9) offers a strong *prima facie* case for the identification of evidence with

[1] Neta (2003) also defends contextualism about evidence ascriptions; I'll contrast my approach with his in §3.13 below.

[2] These first two aims are roughly the same as the project of Ichikawa (2013); the corresponding sections in this chapter are a development of the ideas in that paper.

knowledge, at least once one is focused on the propositional sense of knowledge.[3] According to this approach, a subject's total evidence comprises her total knowledge. So a particular proposition is among a subject's evidence if and only if it is among her knowledge.

In introducing E=K so far, I have been writing in the object language; according to the kind of contextualism I have been defending, sentences like 'a subject's total evidence comprises her total knowledge' will be context-sensitive. However, on the particular contextualist version of E=K I will go on to articulate and endorse, this sentence will come out true in all contexts, because 'evidence' will shift its extension in just the same way 'knowledge' does. For now, however, in introducing the issues surrounding E=K, I propose to set contextualism aside. This will allow me to (a) follow Williamson and his critics more closely, since they are not contextualists, and (b) to demonstrate a point at which introducing resources distinctive to contextualism can help the defender of E=K. So I will write in the object language for the next few sections.

I don't think the main reason to adopt E=K is that there is any one clear argument that establishes it.[4] In general, it seems to me that many of the critics who complain that attempted arguments are inconclusive are correct.[5] Rather, I think that E=K performs well with respect to the intuitive extension of evidence, and has other theoretical advantages, such as being a simple and parsimonious theory. If evidence is propositional, then there must be some kind of story about what it is for a given subject to possess a piece of evidence. Possessing is an epistemic property; knowing is a natural candidate explication of it.

Williamson offers a thought experiment intended to demonstrate that highly probable true belief is not sufficient for evidence:

> Suppose that balls are drawn from a bag, with replacement. ... For a suitable number n, the following situation can arise. I have seen draws 1 to n; each was red (produced a red ball). I have not yet seen draw n+1. I reason probabilistically, and form a justified belief that draw n+1 was red too. My belief is in fact true. But I do not know that draw n+1 was red. Consider two false hypotheses:
>
> > h: Draws 1 to n were red; draw n+1 was black.
> > h*: Draw 1 was black; draws 2 to n+1 were red.
>
> It is natural to say that h is consistent with my evidence and that is h* not. In particular, it is consistent with my evidence that draw n+1 was black; it is not consistent with my evidence that draw 1 was black. Thus my evidence does not include the proposition that draw n+1 was red. Why not? After all, by hypothesis I have a justified true belief that it was red. The obvious answer is that I do not *know* that draw n+1 was red; the unsatisfied necessary condition for evidence is knowledge.[6]

[3] As many writers (including Williamson himself) have emphasised, there are non-propositional ways of discussing evidence as well—for example, there is sometimes evidence in evidence lockers, but the latter presumably do not contain propositions. See Williamson (2000, p. 194). (Compare also authors like Markosian (2014, p. 166), who holds that sensations, a priori intuitions, and seemings are the three main kinds of evidence. But Ichikawa and Jarvis (2013, chs. 12–13) argue that a priori intuitions are not in any interesting sense evidence.) Nevertheless, it seems clear that there is at least *a* sense of 'evidence' that is propositional, and it seems plausible to treat it as the fundamental one for epistemic purposes. Neta (2008b, pp. 95–100) dissents from Williamson's arguments, but endorses the propositionality of evidence on different grounds.

[4] Williamson himself *does* give a simple and valid argument for E=K which seems to me sound: 'All evidence is propositional. All propositional evidence is knowledge. All knowledge is evidence. Therefore, all and only knowledge is evidence.' Williamson (2000, p. 193) But (as Williamson observes) this argument is not cogent; it's not intended to convince dissenters.

[5] E.g. Goldman (2009), McGlynn (2014, p. 81).

[6] Williamson (2000, pp. 200–1)

Although I think that this case does put pressure in favour of knowledge being necessary for evidence, I find Williamson's discussion unsatisfactory in at least two respects. For one thing, as I will suggest in Ch. 4, it seems to me dubious whether one would have a justified belief in Williamson's case.[7] For another, as Brian Weatherson has argued, it may appear as if reasoning parallel to Williamson's own would demonstrate that knowledge itself is insufficient for evidence. I turn to Weatherson's argument now.

3.2 An Argument Against E=K

Here's an argument that Brian Weatherson has proposed against E=K.[8] To set up the argument, take an instance of *inductive knowledge*, defined as knowledge on the basis of evidence that does not entail it. So let S know that p on the basis of e, where $e \,\&\, \sim p$ is possible.

Now consider these two false hypotheses:

(1) $e \,\&\, \sim p$

(2) $\sim e \,\&\, p$

Weatherson's intuition is that under the circumstances, S's evidence is consistent with (1), but inconsistent with (2). But by hypothesis, S knows p; so S's knowledge is inconsistent with both (1) and (2). Insofar as it is intuitive that S's evidence is consistent with (1), E=K seems to get the wrong result; not all of S's knowledge is evidence.

Take a case. Suppose the Usher has the following evidence:

e: The judge has a particular dreamy look in his eye as he watches the plaintiff

On the basis of an abductive inference from this evidence, he comes to know:

h: The judge is attracted to the plaintiff.

It's critical for the case that the usher comes to *know* h, even though it is *possible* for the judge to have the dreamy look in question without attraction. Weatherson's argument would have it that that possibility—the possibility that he has the dreamy look but no attraction—is consistent with the usher's evidence; but it is obviously not consistent with the known fact that the judge is attracted to the plaintiff.

One way to resist Weatherson's argument would be to exploit the context-sensitivity of 'the usher's evidence'. One might adopt a kind of context-sensitivity about the word 'evidence' itself—I'll develop such a line in the sections that follow. But in fact, everyone should admit that there is *some* potential context-sensitivity at play here, even without invoking shiftiness with respect to the word 'evidence'—run-of-the-mill quantifier domain restriction also has a role to play. The defender of E=K can admit with Weatherson that there is a certain reading according to which 'the usher's evidence is consistent with the possibility that the judge has a dreamy look but no attraction' is true—it is true relative to contexts in which the fact that he has attraction, though part of the usher's evidence, is not among the evidence being quantified over. In the Lewisian language of my Ch.

[7] In fact, Williamson himself would presumably now deny that there is justified belief in this case, given some of his later commitments. See for example Williamson (2016) for his commitment to the idea that justified belief itself requires knowledge. See also §4.3.

[8] Weatherson (2009). I also saw, at a 'Basic Knowledge' workshop at the Northern Institute of Philosophy in Aberdeen in 2013, a presentation by Elia Zardini giving a very similar (independently developed) argument. What I have to say about Weatherson's argument applies equally to Zardini's.

1, that is evidence that is 'properly ignored'. This is not a mere *ad hoc* move; after all, a particular bit of evidence was just mentioned and raised to salience: that the judge has a dreamy look. It shouldn't be surprising for it to be easy for talk of 'the usher's evidence' to be restricted accordingly. But unless the clear intuition is that *the totality* of S's evidence is consistent with (1), there is no counterexample to E=K.

It may be helpful to consider an analogy. Suppose Edwin alleges that he likes to smoke, and I go on to observe that 'Edwin's testimony is consistent with the possibility that he didn't smoke yesterday.' This has a true reading even if yesterday he also testified that he'd been smoking that day. My utterance is naturally read as being only about the allegation just mentioned; it does not imply that the *totality* of allegations Edwin has made are consistent with this proposition. In the same way, when Weatherson cites the intuition that 'S's evidence' is consistent with the falsity of the inductively known hypothesis, it is no threat to E=K if we can make it plausible that it should be interpreted as being only about that evidence just mentioned.

So Weatherson's argument may depend on a potentially contentious interpretation of the intuitions in question. Nevertheless, let's grant, if only for the purpose of argument, that it is intuitive that the *totality of* the usher's evidence is consistent with the possibility that the judge has a dreamy look but no attraction. As we will see, there are other ways to resist the argument against E=K.

As mentioned above, the argument in question is quite similar to Williamson's own argument against the sufficiency of probabilistically supported belief for evidence. In that case, one had an inductively well-confirmed belief that was not knowledge, and Williamson cited the intuition that it was not among the evidence. But Williamson's example can be converted to an instance of Weatherson's schema, if we suppose that one *does* end up with inductive knowledge.[9] Suppose for instance that, in a case like Williamson's ball-drawing case, there is a bag with only ten balls in it, and one makes one hundred draws with replacement, and finds each to be red. On this basis, one might well conclude, in an apparent case of inductive knowledge, that the next draw will also be red. But perhaps it is intuitive that the totality of one's evidence is nevertheless consistent with the next ball's being black. (If you're not convinced that 100 draws is enough for knowledge, one can raise the count arbitrarily high.) In such a case, a Weathersonian intuition might have it that (3) is inconsistent with my evidence, while (4) is merely improbable, conditional on my evidence:

(3) Some of the first hundred draws were black, and ball 101 will be red.

(4) The first hundred draws were red, and the ball 101 will be black.

Again, we have an apparent bit of knowledge—that ball 101 will be red—that is inconsistent with (4); but (4) is (I grant for the purpose of argument) intuitively consistent with the totality of the evidence. If so, then E=K appears to have a problem. This, in a nutshell, is Weatherson's argument.[10]

3.3 Evidence as Non-Inferential?

In light of the Williamsonian arguments for the *necessity* of knowledge for evidence, and of Weatherson's set of apparent counterexamples to its *sufficiency*, one natural possibility would be to suppose that something *stronger* than knowledge is required for evidence. Perhaps, for instance, evidence

[9] This point is related to that of McGlynn (2014, p. 54), who argues that Williamson is wrong to suppose, in the most straightforward versions of his case, that one will have justified true belief without knowledge.

[10] In his paper, Weatherson actually gives two arguments against E=K: the one discussed here against the sufficiency of knowledge for evidence, as well as another against the necessity. That latter argument motivates a non-factive conception of evidence. For arguments that non-factive conceptions of evidence are untenable, see Williamson (2000, p. 201), Littlejohn (2012, p. 101), or Littlejohn (2013).

should be identified with *non-inferential* knowledge. While this modification would avoid the objection, it comes with costs of its own. For often, it is intuitive that particular contents which are only known inferentially have a status as evidence. For example, I don't know whether Alan likes whisky, but I do have some evidence that suggests that he does. In particular, that he is Scottish is some, albeit inconclusive, evidence that Alan likes whisky. This is true even though it is not very plausible that I know *non-inferentially* that Alan is Scottish; this knowledge is based on my knowledge of his accent, and memory of conversations I've had.[11] To foreshadow some of the motivation for contextualist approach I'll be developing below, it seems that ascriptions of evidence are at least sometimes lax enough to include indirect knowledge.

Another kind of response in a similar spirit is the approach to evidence offered in Alvin Goldman's (2009) discussion of *Knowledge and Its Limits*. There, Goldman suggests that evidence should be identified with those propositions for which one has *non-inferential propositional justification*. This is, like the proposal just mentioned, stronger than E=K in demanding a non-inferential structure of warrant (and so it comes with the costs just mentioned); but it is also weaker than E=K in some ways, since non-inferential propositional justification doesn't entail knowledge. One implication of this approach is that evidence will not be factive, on the assumption that falsehoods are sometimes non-inferentially justified. As I indicated in the previous section, I find this consequence unacceptable. But there is a subtler drawback to Goldman's approach as well: it closes off the prima facie plausible option of explaining propositional justification in terms of evidence. A promising characterisation of propositional justification is that a proposition is justified for a subject if and only if that subject's evidence supports believing that proposition.[12] But it is difficult to see how one could embrace both of these kinds of views, on pain of circularity.

3.4 Alexander Bird and 'Holmesian Inference'

One assumption of Weatherson's argument is that there is such a thing as abductive knowledge—knowledge that is not entailed by the evidence upon which it is based. One might think that such an assumption is mandatory for any non-skeptic, but this would be a mistake. We should be open to the possibility that there is no abductive knowledge, not because the things we take ourselves to know abductively fail to amount to knowledge, but because many of the things we take ourselves to know abductively are actually based on entailing evidence. One particularly nice implementation of this thought is the treatment Alexander Bird (2005) offers for inference to the best explanation. As Bird emphasises, if one adopts E=K, one has considerably more resources for using evidence to rule out possibilities than one might initially have thought.

For example, consider again the balls drawn with replacement (the knowledge-involving version discussed in §3.2). There, one ends up with knowledge that the 101st draw will be red. According to Bird, the typical way to arrive at this kind of knowledge is via a kind of 'Holmesian Inference' (after Sherlock Holmes's famous endorsement of deduction: 'once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth'). If I know that there are ten balls in the bag, and that I have drawn and replaced balls 100 times, and that each time I observed a red ball, then I am in a position to choose among competing hypotheses. Here are some possibilities:

h_1. Every ball in the bag is red.

h_2. There is at least one black ball, which I've coincidentally never drawn.

[11] As a matter of fact, I haven't even heard that Alan is Scottish first-hand, i.e., from Alan himself.

[12] See for instance Ichikawa and Jarvis (2013, pp. 162) for one such characterisation. But note that Goldman himself isn't likely to be interested in such an evidentialism; see Goldman (1986, pp. 89–93).

> h_3. There is at least one black ball, which has a magical ability to avoid my fingers when I reach in.
>
> h_4. etc.

It's difficult to deny, without accepting a radically general form of skepticism, that somehow or other I may end up with knowledge of h_1. Bird's observation is that this knowledge may well depend in essential part on the knowledge that h_2, h_3, etc. are false. For example, one may eliminate h_2 on the known grounds that if there were a black ball in there and no explanation for why I shouldn't draw it, I would have drawn it by now. One may elimiate h_3 on the known grounds that there is no such thing as magic balls. Etc. The relevant knowledge and process of elimination here is typically tacit, but that doesn't stop it from being known and therefore part of the evidence, according to E=K.

If something like this picture is right, then Weatherson makes problematic stipulations about his schema for counterexamples to E=K. For it is essential to Weatherson's case that we're looking at cases of knowledge that are not entailed by the evidence upon which they're based. If Bird's implementation of the E=K idea is right, then there are no such cases. One isn't concluding that the ball will be red *merely* on the basis of the past draws; part of the basis is the known fact that every ball is red. Or, to return to the case of the usher and the judge, the former doesn't know that the latter is attracted to the plaintiff *merely* on the basis of the look in his eye; part of the basis is the known fact that he wouldn't have that look if he weren't attracted to her.

So insofar as Weatherson means to be giving an *objection* to E=K, he may well be begging important questions. If Bird's approach to abduction is the right one, then there are no cases of knowledge based on inconclusive evidence (even though there's probably about as much knowledge as we think there is).

But let's set Bird's approach aside. As we will see, interesting possibilities emerge even if we grant Weatherson the assumption that there is genuinely abductive knowledge.

3.5 Abominable Conjunctions

I mean from here on to grant both of the following to Weatherson for the purpose of argument:

- There are cases of genuinely abductive knowledge, where S knows p on the basis of e, even though e does not entail p.
- In such cases, it is intuitive that the totality of S's evidence is consistent with the possibility of not-p.

Weatherson argues that if we accept the latter intuition, we will have to abandon E=K; p is stipulated to be known, but is intuitively consistent with not-e. So if the evidence includes all knowledge, it must include p, and so couldn't be consistent with not-e. We should tread carefully around arguments like this one; Weatherson's argument is structurally very parallel to a particularly controversial treatment of the skeptical paradox. The key thing to notice is that the intuition cited in the second bullet point above is in essence a kind of skeptical intuition. Notice that this argument too has *prima facie* force behind it:

> Anti-Closure. There are cases of genuinely abductive knowledge, where S knows p on the basis of e, even though e does not entail p. In such cases, it is intuitive that S is not in a position to know that $(e \ \& \sim p)$ is false. So in general knowledge cannot be closed under logical deduction.

For example, suppose one knows that an animal is a zebra on the (inconclusive) basis of its appearance; it is intuitive that one does *not* know that it isn't a cleverly disguised mule. I discussed attempts to deny closure in this kind of radical way in Ch. 2; Nozick (1981) and Dretske (1981) were the prominent defenders of this kind of approach. Notoriously, however, this kind of line on closure implies the truth of abominable conjunctions—in this case, 'S knows that *p*, but doesn't know whether (*e* & ∼*p*).' Most epistemologists find this result unacceptable, and so most epistemologists are committed to resisting the anti-closure argument.[13] See §2.1. But given the parallel structure of Weatherson's argument against E=K and the 'abominabilist' argument against closure, it is worth pausing to consider to what degree the standard responses to the latter extend to Weatherson's arguments.

For concreteness, let's take an example of an argument for an abominable conjunction: Angelina says she bought a dress, and Edwin believes her. She's speaking truthfully, and Edwin has no special reason to doubt her testimony, so the 'ordinary' knowledge claim is intuitively true:

> o. Edwin knows that Angelina bought a dress today.

But it is *possible* that Angelina might have been speaking falsely when she said she'd bought a dress, and that's the kind of 'skeptical' scenario that intuitively, Edwin isn't able to discern.

> s. Edwin doesn't know that Angelina wasn't lying when she said she bought a dress today.

But the conjunction of these intuitive verdicts is abominable:

> o & s. Edwin knows that Angelina bought a dress today, but he doesn't know that Angelina wasn't lying when she said she bought a dress.

Cases like these are familiar from the discussion of Ch. 2.

The closure-denying argument, like Weatherson's E=K-denying argument, has three elements: a claim of ordinary (non-basic) knowledge, an intuition to the effect that some skeptical scenario inconsistent with that ordinary knowledge is consistent with either the subject's knowledge or evidence, and a conclusion conjoining those results that refutes the denied principle. Once it is evident that Weatherson's denial of evidence has so much in common with the closure-denier's denial of knowledge, the idea of offering parallel treatments to the cases seems rather attractive.

I have already signalled my allegiance to a contextualist response to the closure-denier; accordingly, I will focus below on a parallel contextualist response to Weatherson's argument against E=K. First, however, I'll canvass other, non-contextualist, approaches in §3.6 below, before developing my preferred contextualist interpretation in §3.7.

[13] This is presumably true even of those epistemologists who want to deny that closure holds *in generality* for reasons other than the kinds articulated here. For example, Maria Lasonen-Aarnio (2008) has argued that a safety requirement on knowledge may imply that in certain special cases, single-premise deduction can fail to extend knowledge, due to the deductive risk associated with inference. Her idea is that since there is some small possibility that one will make a logical error in attempting a deduction, deduction adds risk; if one's initial knowledge is already risky enough to be near the threshold, the extra risk associated with a deduction might prevent the conclusion from counting as known. (She doesn't *endorse* this line, but argues that most deniers of multi-premise closure are committed to it.) Lasonen-Aarnio's suggestion does imply that there are certain special cases of true abominable conjunctions—when one *just barely* knows that *p*, it might be true that one knows *p* but is not in a position to know *(p v q)*, say. Perhaps, for reasons having to do with vagueness, it will be impossible for ordinary people to know which propositions serve as counterexamples. But her view does not predict *bald* counterexamples to single-premise closure, i.e., cases of clear, non-marginal knowledge conjoined with ignorance of trivial logical consequences of that knowledge. And the skeptical intuitions cited in the main text here are not limited to the margins. So even moderate deniers of single-premise can and should take steps to avoid these unpalatable consequences. See 2.6, fn. **??**.

3.6 Non-Contextualist Responses

Other than contextualism, there are, it seems to me, three main ways of avoiding the argument for abominable conjunctions: skepticism, Mooreanism, and shifty invariantism. Each has an analogue in the evidential argument that is at least as plausible as it is in this context. I'll consider the three in turn.

Skepticism is the simplest option. It denies the intuitive verdict that one possesses ordinary knowledge in the cases in question. This straightforwardly avoids abominable conjunctions, which begin with an ascription of ordinary knowledge; it also avoids Weatherson's argument against E=K, since it has it that the intuitive cases of abductive knowledge aren't genuine knowledge. So, by E=K, they aren't evidence; so the intuitive verdict that the totality of the evidence is consistent with the falsity of the knowledge is vindicated.

The other straightforward response to the argument for abominable conjunctions is to reject the skeptical intuition, perhaps on the basis of the ordinary knowledge intuition. This 'Moorean' response to skepticism enjoys wide popular support among epistemologists.[14] On this view, Edwin *does* know that Angelina isn't lying about buying a dress, even though he hasn't performed a test that would detect such a lie. (So this approach is inconsistent with a sensitivity requirement for knowledge; see Ch. 2.) One of the natural outstanding questions for this approach, then, concerns how it is that one has this kind of insensitive knowledge. Another challenge for this kind of approach is to explain the source of the skeptical intuitions; if Edwin really does know that Angelina isn't lying, why isn't that fact more evident to us as we think about his epistemic position? The skeptical intuitions don't derive from recognition of truth; so where do they come from? Given a fuller implementation of the Moorean strategy, it seems to me very plausible that the resources used to explain away skeptical intuitions about knowledge may generalise just as well to explain away skeptical intuitions to the effect that the subject's evidence is consistent with the skeptical scenario. For example, Lackey (2008, pp. 137–8) argues that we're inclined mistakenly to intuit that we don't know that our lottery tickets will lose because asserting that we knew such a thing would falsely implicate that we have distinctive, non-public evidence to that effect.[15] If anything like this story could be made to work, it would also presumably explain why we mistakenly suppose that our evidence doesn't rule out the possibility that the ticket will win—asserting that it did might also falsely implicate that the speaker is in possession of special evidence that is not in the common ground.[16] In short, Moorean defenders of ordinary knowledge, who also readily attribute knowledge of the negations of skeptical scenarios, need to have the resources to deny certain skeptical intuitions, and to posit non-obvious explanations of positive epistemic statuses. There is every reason to suppose that these resources could also be used to deny the skeptical intuition that is central in Weatherson's case against E=K.

The final non-contextualist response to the abominable conjunctions argument I'll consider is the kind of 'shifty invariantism' associated with the names 'interest-relative invariantism' or 'subject-

[14] See for example Sosa (1999), Pritchard (2007), Brown (2005).

[15] See also Douven (2012) for an endorsement of a similar line.

[16] In fact, I don't think this explanation works to explain away the intuition that we don't know the ticket won't win. (This, after all, is part of why I'm a contextualist.) I find this particular invariantist story inadequate because (a) the putative implicature seems ad hoc, not derived from general conversational principles—see Dimmock and Huvenes (2014) and DeRose (2002) for arguments of this form applied to similar strategies about different cases—and (b) the explanation seems to overgeneralise, predicting that we should also intuit that we don't even know that it's very probable that the ticket will lose—see Williamson (2000, p. 247); see also Benton (2014, n. 28). My point here is the comparative or conditional one: the story works *as well* in the evidence case as in the knowledge case; or *if* the story works in the knowledge case, it works in the evidence case as well. To fully decide this question, one would want to articulate the best version of the explanation in the knowledge case, then apply it to the evidence case. This is well beyond my present scope, in part because I am not satisfied by any version of the story in the knowledge case, but I hope that I have made it plausible that the prospects for theories of these kinds may rise and fall together.

sensitive invariantism'.[17] According to the kind of view I have in mind, a subject might shift from a state of knowledge to a state of ignorance, depending on whether she is thinking about a given possibility.[18] On this kind of view, whether Edwin knows that Angelina bought a dress constitutively depends on whether he's thinking about the possibility that she might be lying; if he is ignoring the possibility, he has knowledge, but if it has entered into his mind, he doesn't. Consequently, on such a view, E=K predicts that what *evidence* one has *also* depends constitutively on what possibilities one is considering. While he is thinking about the possibility that she is lying, his evidence does not include the proposition that she bought a dress; it might instead be restricted to something weaker, like the proposition that she said she bought a dress. But if he is ignoring the possibility, his evidence does include the proposition that she bought a dress (and so his evidence is inconsistent with the ignored possibility). Or, to return to the case of the red balls discussed above, one might possess the knowledge that the 101st draw will be red only if one isn't entertaining the possibility that there is a black ball that has eluded capture 100 times.

The degree to which this kind of approach avoids the Weathersonian objection is debatable. Certainly, if our subject is thinking about the skeptical possibility in question, he won't have knowledge, and so, by E=K, he won't have the intuitively problematic piece of evidence. In this version of the case, the shifty invariantist offers the same treatment as the skeptic does. But what if the subject is ignoring the skeptical possibility? Suppose the thought that Angelina could be lying never enters Edwin's mind; she tells her that she bought a dress, and he believes her straightaway, without considering the likelihood that she speaks truly. (This is plausibly the kind of 'default' way most of us accept testimony most of the time.) It is an implication of E=K, on this approach, that if Edwin is ignoring skeptical possibilities, then the proposition that Angelina bought a dress *is* among his evidence. Does this respect the skeptical Weathersonian intuition? At first glance it appears not: that intuition was an intuition to the effect that the totality of Edwin's evidence is consistent with the possibility that she was lying, and didn't buy a dress. But this clash may be mollified by two observations.

First, *were Edwin to consider the question* whether the totality of his evidence was consistent with the possibility that she was lying and hand't bought a dress, the shifty invariantist interpretation of E=K has it that the answer would be no. Edwin's evidence against this possibility is *elusive* in the sense that, were he to consider whether he has such evidence, it would disappear.[19] Perhaps this is an adequate way of respecting the intuition that the evidence is consistent with the skeptical hypothesis.

---

[17] I have here in mind e.g. the views of Hawthorne (2004), Stanley (2005) and Weatherson (2005). Note that the 'impurism' of Fantl and McGrath (2009) is *not* an instance of shifty invariantism; although Fantl and McGrath embrace a kind of pragmatic encroachment, they are neutral on the question of invariantism. As I indicated in §1.9, the extant names for this family of views are all problematic in various respects. (It is less than obvious to me that there is anything centrally unifying to these views, other than the historical context in which they were developed.)

[18] This characterisation is incomplete; all non-skeptics, not just shifty invariantists, should admit the possibility that thinking about a possibility causes the subject to lose confidence to a degree amounting to a loss of knowledge. This is a core element in Jennifer Nagel's (2010) defence of non-shifty, non-skeptical invariantism. Categorisation of the views becomes even more complicated if, as Clarke (2013) has argued, thinking about skeptical possibilities can *constitutively* (as opposed to merely causally) make for a difference in belief.

[19] My use of the term is inspired by David Lewis's use in 'Elusive Knowledge', although Lewis's use is misleading in certain respects. At times, Lewis writes as if he is defending the kind of shifty invariantism discussed here, as when he writes that 'it will be inevitable that epistemology must destroy knowledge. That is how knowledge is elusive. Examine it, and straightway it vanishes.' Lewis (1996, p. 560). But it doesn't follow from the contextualist view he defends that examining knowledge makes it vanish; one doesn't lose knowledge by considering skeptical possibilities—one merely moves into conversational contexts in which 'knows' no longer applies. (Compare: if I am pointing to her, and then I move my arm and point to him instead, I didn't make her or anyone vanish; I just made her stop being the referent of 'her'. Moving my arm around isn't, except in very special circumstances, a way to destroy her or anyone else.) But shifty invariantism *is* a view that satisfies Lewis's description, whereupon when one reflects on skeptical scenarios and the inconclusive evidence for our beliefs, this mental act destroys knowledge.

Second, the intuition that Edwin's evidence is consistent with the skeptical hypothesis in question is itself somewhat shifty. (Unsurprisingly, the contextualist treatment I'll develop in the next section will exploit this fact.) One way to reconcile skeptical intuitions with the shifty invariantist approach to E=K is to focus on *other* hypotheses for which the abductive knowledge might be evidence. Suppose again that the shifty invariantist view is right, and that Edwin is ignoring the far-fetched possibility that Angelina is lying about the dress. So he knows, albeit abductively, that Angelina bought a dress today. Now suppose Edwin is asked to consider another hypothesis. Suppose his friend, who is concerned about the recent downturn in the fashion industry, speculates that *no one has sold any dresses in the past week*. Under the circumstances, it seems to me very plausible that Edwin has conclusive evidence against this hypothesis—in particular, that *Angelina bought a dress today*. If this is right, then in the circumstances mentioned, Edwin *does* have evidence that is inconsistent with the possibility that Angela was lying about having bought a dress. So it seems that the shifty invariantist approach to E=K can credibly accommodate at least a significant amount of the data.

Readers familiar with the literature on contextualism and shifty invariantism will naturally wonder about third-person ascriptions of evidence, where the speaker's interests and attention come apart from those of the subject. As in the parallel discussion of 'knows', shifty invariantism performs less well with respect to these cases. Suppose that we—you the reader, and I the author—do *not* ignore the possibility that Angelina is lying, but that Edwin (who isn't reading this) does. Is the totality of Edwin's evidence consistent with the possibility that Angelina is lying? The Weathersonian skeptical intuition—that it does—seems to me to persist. But the shifty invariantist E=K theorist cannot yield that verdict. On that view, only Edwin's own attention or interests can make for differences in Edwin's evidence; you and I have no effect. As I'll go on, predictably enough, to argue in the next section, the contextualist is in a better position here. But first it is proper to spend a few more words in defence of the shifty invariantist approach.

Notice that the challenge for this shifty invariantist approach to evidence is exactly analogous to a parallel challenge in the case of knowledge. It is widely recognised that so-called 'interest-relative invariantism' delivers counterintuitive results with respect to knowledge ascriptions by speakers in high-standards contexts about subjects in low-standards contexts.[20] Shifty invariantists are well aware of this feature of their view; many of them have developed attempts to explain away the relevant intuition.[21] As it happens, I'm not particularly impressed by the explanations on offer;[22] but if one *is* sympathetic to the kind of line advanced about such third-person knowledge ascriptions, there looks, as in the case of the Moorean dialectic mentioned above, to be every reason to expect it to extend to the problematic cases of third-person *evidence* ascriptions too. So the general strategy of this section seems at least promising. Everybody faces a challenge from the threat of abominable conjunctions; each non-contextualist response in this knowledge case may well generalise against Weatherson's argument against E=K.

[20] See for instance the 'High-Attributor, Low-Subject' case discussed in Stanley (2005, pp. 5–7). Note that it is for just this reason that Fantl and McGrath (2009) opt for a contextualist interpretation of their 'impurism', which is motivated in ways very similar to the motivations of the shifty invariantists. Defenders of shifty invariantism often claim parity with contextualists on this score, citing skeptical intuitions about 'Low-Attributor, High-Subject' cases, which, they say, their view delivers and contextualism does not. See e.g. Stanley (2005, p. 116). I agree with DeRose (2009, pp. 238–41) that contextualists are able to capture this data without especial difficulty.

[21] E.g. Stanley (2005, pp. 101–03), Hawthorne (2004, pp. 164–5). See also Nagel (2010, pp. 424–5), who defends a similar strategy for a parallel challenge to her view; however, Nagel does not consider her view to be in the relevant respect a kind of shifty invariantism.

[22] Ichikawa (2009a, pp. 102–4) attempts to rebut the particular strategy of Stanley (2005).

3.7 Contextualist E=K

Broadly speaking, there are three ways for contextualists about 'knows' to endorse E=K.

One way is to identify evidence with the propositions that can be truly described as 'known' in some particular, favoured context. Consider for example the particular context Timothy Williamson was in when he wrote Chapter 9 of *Knowledge and Its Limits*. A contextualist of this first sort will think that the sentences he wrote in that context, e.g. 'All propositional evidence is knowledge'[23] are context-sensitive, and assent simply to the proposition expressed by them in that particular context. Suppose for simplicity an indexical model of the context-sensitivity of 'knows', and that in the context in question, 'knowledge' referred to property K_i. This property may be identified with evidence. The contextualist of this first sort is a contextualist about 'knows' only—she is an invariantist about 'evidence'. In particular, she thinks the latter always refers to K_i.

The contextualist who takes this tack will admit that there are some conversational contexts relative to which 'E=K' expresses a falsehood—namely, any in which 'K' refers to a property other than K_i. There's nothing incoherent or obviously unacceptable about such a metalinguistically specific identity claim; compare this truth: I am Jonathan. This is a plausible claim that is both informative and metaphysically necessary, like E=K is meant to be; but there is no pressure, from accepting it, to accepting that 'I am Jonathan' will be true in all contexts—it manifestly isn't. (Try saying it yourself.) There's no reason a contextualist couldn't think of E=K in a similar sort of way.

If you go this way, then you can accommodate the kind of example we're worrying about by suggesting that evidence is identified with a rather stringent 'knows' relation—one that does not apply to the apparent cases of abductive knowledge. Probably the best way to hold this view would be to say that it does not apply to any abductive beliefs. Two obvious candidates for epistemic standards for 'knows' come to mind:

1. A standard that countenances only basic beliefs (beliefs that are not based on any other beliefs)

2. A standard that countenances only basic beliefs and beliefs produced via deduction from countenanced beliefs

The idea would be that the special, privileged interpretation of 'E=K' will be the one produced in a context with standards like these. (This is *not* the view that these are the only standards for 'knows'—it's just the view that these are the special standards relative to which 'E=K' expresses an important epistemological truth.) For the purpose of accommodating Weatherson's intuition that the evidence is *consistent* with the abductive 'knowledge', either of these standards will suffice. They differ only with respect to whether non-basic beliefs deductively entailed by evidence literally are evidence.

The contextualist of this first type can agree with Weatherson that abductive knowledge isn't evidence. But relative to the context in which 'E=K' expresses a truth, such conclusions don't count as 'known' either. This is not, however, to capitulate to extreme skepticism, the way Weatherson suggests it is, since it insists that many utterances (in lower-standard contexts) of 'S knows p' can be true when S's belief that p is based on non-entailing evidence. One will not, however, be able to use this kind of contextualism to explain the shifty patterns in evidence ascriptions mentioned above.

A second way a contextualist might interpret 'E=K' would be to endorse a kind of subject-centred metalinguistic generalisation of it, such as this one:

> S has p as evidence if and only if 'S knows p' is true relative to S's context.

[23] Williamson (2000, p. 193)

Depending on the details of how context interacts with truth conditions, this treatment of E=K could result in a treatment very much like that of the shifty invariantist discussed above. On the Lewisian approach I laid out in Ch. 1, it will predict that a given piece of abductive knowledge is among S's evidence only when S is ignoring certain not-S probabilities.[24] Note that this view will countenance some statements that are in superficial tension with a commitment to E=K—for example, there will be contexts relative to which these sentences express truths:

- 'S knows *p* but *p* isn't part of S's evidence.'
- '*p* is part of S's evidence, but S does not know *p*.'

(The contextualist of the first type will also accept these as true in many contexts.) I don't see that the contextualist E=K theorist need be embarrassed by these results; it's not like the statement of E=K sounds obviously tautological. But it is worth being explicit about these being commitments. John Hawthorne considers a parallel observation as an objection to certain contextualist implementations of some other principles regarding knowledge.[25] I'll return to those arguments in Chs. 5 and 6.

A more serious problem for this approach, it seems to me, is the one indicated in the previous section: this kind of view has difficulty predicting skeptical intuitions about evidence ascriptions made by speakers in high-standard contexts about subjects in low-standard contexts. (But if there are good answers to those worries, then the contextualist of this persuasion may presumably be able to make use of them as well.)

A third contextualist implementation would have 'E=K' express a truth in *all* contexts. On this implementation, sentences like 'S knows *p* but *p* isn't part of S's evidence' will *never* express truths. Since by hypothesis, we're considering a kind of contextualism about 'knows', this kind of contextualist will also have to embrace a parallel contextualism about 'evidence'.[26] On this view, any context in which 'S knows *p*' is true will also be one in which '*p* is among S's total evidence' is true. The conversational context will provide the epistemic standard required for a subject to count as having 'evidence', just as it does for 'knowledge'—and indeed, it will be the same standard. (The view is closely analogous to the relationship between knowledge ascriptions and counterfactual conditionals described in Ch. 2.)

The contextualist who takes this approach can make all the same moves as the IRI theorist did about first-person knowledge ascriptions and evidence ascriptions; he also makes parallel moves for third-person cases. If I'm in a more skeptical context, I'll say 'Edwin doesn't know that Angelina bought a dress today' and speak truly, and I'll also say 'Edwin's total evidence is consistent with the possibility that Angelina was lying about buying a dress' and speak truly; but in other, less skeptical contexts, I could truly affirm both 'knowledge' and 'evidence'. One can think of this (as with the relationship between contextualism about 'knows' and shifty invariantism) as a kind of metasemantic recasting of the shifty invariantist solution. But crucially, this contextualist can also extend the treatment to third-person cases: if I am in a skeptical context, focused on the mere abductive

[24] The not-S possibilities in question are those 'uneliminated by S's evidence'; so the proper understanding of evidence will complicate the view further. I'll discuss this at length in §3.10 below.

[25] Hawthorne (2004, pp. 88–9)

[26] I assume there is no plausible context-sensitivity in the '=' that could be used to avoid this commitment. Note that if a 'knows' contextualist only endorsed the *sufficiency* of knowledge for evidence, one could hold a statement of the relationship between 'knows' and 'evidence' in metalinguistic generality without committing to contextualism about 'evidence' by supposing that satisfying the weakest possible standard for 'knows' (assuming there is such a thing) suffices for evidence. The same goes *mutatis mutandis* for the necessity of knowledge for evidence and the strongest possible standard. But I can't see how a contextualist about 'knows' who endorses both directions of E=K can hold the latter in metalinguistic generality without contextualism about 'evidence'.

grounds for Edwin's belief, but Edwin himself is taking its truth for granted, then, according to this contextualist approach, I can deny that Edwin has the claim in question as 'evidence' just as I deny that he has 'knowledge'. As we have seen, the shifty invariantist's failure to deliver this intuitive verdict is one of its main shortcomings.

I take this to be reason to prefer this third kind of contextualist treatment.[27] The remainder of this chapter will pursue it in more detail.

3.8 The Intuitions Again

I am defending a contextualist implementation of E=K in metalinguistic generality; all knowledge is evidence, and all evidence is knowledge, and 'knows' and 'evidence' are each context-sensitive in a parallel way. But am I respecting the data? We began with a situation in which 'Edwin knows that Angelina bought a dress' was intuitively true, but where 'Edwin's evidence is consistent with the possibility that Angelina was lying about the dress' was also intuitively true. How can there be such situations, if all and only that which is 'evidence' in a context is 'knowledge' in that context?

My view is that 'knowledge' and 'evidence' have the same extension in all contexts. This does not, however, commit us to the claim that they have the same pragmatic profiles. In particular, their use may tend to alter the context in divergent ways. Take some context c_i, such that 'Edwin knows that Angelina bought a dress' and 'Edwin's evidence is inconsistent with the possibility that Angelina was lying about the dress' are both true relative to, but unsaid in, c_i. There is no guarantee that c_k, the context that emerges from an assertion 'S knows p,' made against the background of c_i, will be the same as c_e, the context that would result from an assertion, in c_i, of 'Edwin's evidence is inconsistent with the possibility that Angelina was lying about the dress'. It may be that c_e is more skeptical than is c_k with respect to p. That this situation is possible shows that the disanalogy with respect to the intuitiveness of knowledge ascriptions and evidence ascriptions is consistent with 'E=K' being true in all contexts.

One possible explanation for this sort of effect would be that the word 'evidence' has a more skeptical pragmatic profile than does the word 'knows'. I think there is something to this idea; discussion of evidence is typically a *skeptical* activity; to consider what evidence there is for p is already to take a searching, challenging status vis-à-vis p. We don't typically ask ourselves what evidence we have for those things we are taking for granted; when we do consider that question, we stop taking them for granted, questioning what they rest upon. According to traditional skeptical arguments, we must always do this for all of our beliefs; nonskeptical foundationalists think it is legitimate to stop at some point. According to the view in question, context determines at what point it is permissible to stop, consistent with being said truly to 'know'. To question the evidence for p is often thereby to deny that p is a legitimate stopping point; the context, then, shifts to accommodate this skeptical requirement, and p no longer counts as 'known'. So considering whether something is 'evidence' can have skeptical effects on the conversational score.[28] Considering whether something

[27] But as I mentioned in the previous section, any E=K theorist who's already a shifty invariantist about 'knows' will here find little new pressure towards contextualism, since this point about third-person evidence ascriptions is exactly parallel to the shifty theorist's already-awkward handling of third-person knowledge ascriptions; skeptical 'knowledge'-denying intuitions persist when the attributor, but not the subject, is in a skeptical situation. So the shifty invariantist, like the nonskeptical classical invariantist, has to explain away intuitions about such cases. That the same pattern of intuitions applies to third-person 'evidence' denials is no new problem for the IRI theorist; it's just what the combination of E=K and IRI predicts. If you're an IRI theorist who wants to explain away third-person knowledge denials, it is reasonable to hope that your explanation will extend to evidence denials as well.

[28] Jennifer Nagel's work on 'epistemic anxiety'—e.g. (Nagel, 2010)—may lend some support to this thought. (Though Nagel herself would certainly not endorse this contextualist line.) As Nagel emphasises in that paper, seeking out and reflecting on evidence is an activity characteristic of a kind of epistemic anxiety inconsistent without outright belief (and therefore also inconsistent with knowledge).

is 'known' sometimes, but not always, has similar skeptical effects. Sometimes, questioning whether something is known comprises taking a challenging stance towards it; when you assert p, if I ask whether you know p, I strongly suggest that there's reason to think you don't. This contributes towards a more skeptical conversational score. But the word 'know' does not always have this effect—it doesn't, I think, have it as robustly as does 'evidence'. If I ask you whether you know what time it is, this has very much the same conversational effect as does merely asking you what time it is.[29]

(These considerations also bear on another allegedly counterintuitive consequence of E=K, namely, that all known propositions are self-evident. What evidence does Edwin have that Angelina bought a dress? It sounds wrong to cite that *she bought a dress* as evidence; intuitively, one ought to cite her testimony, or perhaps her receipts. Williamson (2000, p. 187) attributes our reluctance to cite the fact of the purchase as evidence as a mere pragmatic effect; but on the view sketched here, one might maintain that, in the relevant context, 'that she bought a dress is evidence for Edwin' is literally false—the context is one in which that proposition is being tested, and being challenged; it is therefore not one in which we should say it is 'known'. (If we did try to say that, we would plausibly be accommodated; the context would shift to a less skeptical one.) Another kind of contextualist resource would be to invoke the kind of domain restriction mentioned in §3.2 above; the proposition in question really is 'known' relative to the context, and so really is 'evidence' relative to the context, but is not among the evidence being quantified over. Compare the question, 'which people in this room are members of your family?'; you will typically not mention yourself. This illustrates the pragmatic importance on the conversational score of what something is evidence *for*. Natural language discourse rarely discusses whether something is evidence for a given subject *tout court*; it is much more natural to ask whether, for S, p is evidence *for H*. And what hypothesis is under consideration can dramatically affect the context.)

Positing a difference in pragmatic profile of 'evidence' and 'knows' is *one* way to accommodate the intuitive data. But there is another, less controversial strategy for explaining the relevant data as well. Again, what needs explaining are pairs of intuitions like these:

- Edwin knows that Angelina bought a dress — Edwin's evidence is consistent with the possibility that Angelina is lying about the dress
- I know that the next ball will be red — My evidence is consistent with the possibility that the next draw will be black
- etc.

Laid out in this way, it is easy to observe that these are not *minimal pairs*. Whether they use the word 'knows' or 'evidence' is not their only, or even their most conspicuous, difference. The evidential element also *mentions skeptical possibilities*. And on most versions of contextualism, including the one I endorsed in Ch. 1, this is exactly the sort of thing that has a tendency to generate more skeptical conversational contexts. So even if one is unconvinced by my suggestion that

[29] This strategy—that of maintaining that two context-sensitive terms are co-referring in all contexts, but have asymmetric affects on the conversational score, has a few other obvious potential applications. It may permit us, for example, to respect the intuitive notion that knowledge requires certainty. (The intuitive status of this equivalence is encoded, for example, in the abominable nature of 'I know that p but I am not certain that p', or even in the Moore-paradoxicality of 'p but I am not certain that p'. See Stanley (2008).) Knowledge requires certainty; indeed, 'knowledge requires certainty' is true in all contexts, but to describe something as 'certain' or 'uncertain' has a more skeptical effect on the conversational score than does 'knowledge' talk. The same may go for the absolute gradable adjectives featured in Unger (1975). Plausibly, what it is to be 'flat' is to be 'perfectly flat'; each is context-sensitive, and in all contexts, each is equivalent, but the two tend to generate divergent contexts. See §6.1.

'evidence' tends to have more skeptical effects on conversational scores than 'knows' does, a contextualist should predict that *these* discussions of evidence should have more skeptical effects than the knowledge ascriptions to which they're compared.

Notice, for instance, that one can generate the relevant skeptical intuitions just in terms of knowledge:

- Edwin knows that Angelina bought a dress — Edwin's *total knowledge* is consistent with the possibility that Angelina is lying about the dress
- I know that the next ball will be red — My *total knowledge* is consistent with the possibility that the next ball will be black

Indeed, these are just the kinds of pairs that motivate the skeptical paradox about knowledge discussed in Ch. 1. I argued there that contextualism is an attractive solution to that puzzle; I think it's an attractive solution to this puzzle too. Certainly it is better than competing views at predicting the intuitions cited.

3.9 Evidence as Important

One complaint I sometimes hear (more often in conversation than in print) is that if one is a contextualist about both 'knows' and 'evidence', one won't have enough 'fixed elements' in one's epistemology. I confess I have some difficulty understanding just what the relevant complaint here is, but I have heard it enough times that I wanted to discuss it at least briefly. One possible interpretation of the worry is that it is psychologically unsettling to have an epistemic theory with so many context-sensitive words in it; perhaps the theorist might have an uncomfortable feeling of epistemic vertigo. So stated, this is no *objection* to the contextualist approach I'm offering; it might merely be an expression of disappointment that we're not working with a more straightforward language.

I suspect that one source of the concern here might derive from the kind of use–mention confusion discussed in the Introduction. (Recall the case of Reginald the confused narcissist, who perceived a tension between, on the one hand, the content of the thought he expresses when he says to himself, 'I am important and special', and, on the other, the semantic observation that 'I' is a context-sensitive indexical that refers to whoever happens to be speaking in a given context.) Once we keep in mind the fact that both brands of contextualism discussed here—contextualism about 'knows' and contextualism about 'evidence'—are *semantic* claims about particular *English words*, I think it's clear that there's no threat from contextualism to supposing that knowledge or evidence is in any way unimportant. If it turned out that words like 'husband' and 'wife' and 'divorce' and 'honeymoon' were context-sensitive, that would pose no threat to the stability of marriage. Why should epistemology be any different?

So stated as a general worry about stability, I do not here see any objection to my approach that I know how to make sense of. However, there are particular approaches to relations between knowledge and evidence that would require some careful navigating, on my view. So far, I have been operating at a relatively abstract level, discussing the raising and lowering of epistemic standards rather vaguely. It is valuable to have some of the discussion at that level, to point out that contextualists of various sorts have the resources to accommodate E=K in the face of Weatherson's challenge. However, it is also worthwhile to look in more detail at the particular brand of Lewisian contextualism I laid out in Ch. 1. As I mentioned above, the contextualist implementation of E=K has special consequences for that approach, as 'evidence' plays a central role in the Lewisian statement of the truth conditions for knowledge ascriptions. In the next several sections, I'll explore that role, beginning with an investigation into potential worries about circularity.

3.10 Circularity and Basic Knowledge

The contextualist approach I articulated in Ch. 1 is based closely on that of Lewis (1996). Lewis's formulation, and mine, are each given in terms of evidence. So it is worth investigating the impact of my contextualist version of E=K on that account.

To recap some of the central thoughts from Chapter 1, Lewis gives us this account of knowledge:

> S knows proposition P iff P holds in every possibility left uneliminated by S's evidence. (551)

Lewis means for this to be true in all contexts; 'knows', then, exhibits context-sensitivity in a way parallel to the context-sensitivity of 'every possibility'. On Lewis's view, there is a class of possibilities that grows and shrinks according to conversational context; our knowledge ascriptions are true when the subject's evidence eliminates all of the members of that class in which the object of knowledge is false. Lewis is neither a contextualist about 'evidence' nor an E=K theorist. His 'evidence' is invariant and his evidence is internalistic. Is Lewis's view consistent with the contextualist approach to E=K? The combination of views would be:

> In all contexts, "S knows *p* iff *p* holds in every possibility uneliminated by S's knowledge" expresses a truth.

This is not inconsistent, but it's all but vacuous—and the respect in which it is not vacuous is the implausible implication that one always knows all the entailments of all one's knowledge. So this straightforward combination of views is an untenable one.[30] But this doesn't mean we can't or shouldn't integrate the views in a subtler way. Recall, after all, that at least *part* of the reason this combination of views is untenable is that, for reasons articulated in Ch. 1, the bald formulation of Lewis's own view is itself untenable. Whatever one's theory of evidence is, Lewis's own view has it that all subjects always know all the entailments of their evidence. Given his particular internalist approach to evidence, this directly and implausibly implies that introspection is infallible and automatic—one always knows everything about what experiences one is having. It also implies that everyone is logically omniscient. Another odd implication of Lewis's account is that knowledge does not entail belief. Lewis himself was bold enough to recognise and embrace this result.[31] Lewis's remarkable ability to embrace counterintuitive consequences notwithstanding, it is my view that we should not accept knowledge without belief.[32] So there are good independent reasons to revise the particular formulation Lewis gives for his brand of contextualism. In Ch. 1 I offered this object-level statement of the truth conditions for 'S knows that *p*':

> S knows that *p* just in case, for some evidence E, (i) S believes that *p* on the basis of E, and (ii) all the E cases are *p* cases.

Recall that I intend this statement to hold with metalinguistic generality, where the context-sensitivity of 'knows' on the left-hand-side runs parallel to the context-sensitive quantifier domain restriction associated with the 'all' on the right-hand-side. This version of the Lewisian view fits much more happily with the contextualist treatment of E=K. Filling this in explicitly, we have:

[30] See the related discussion in §1.7.

[31] 'I even allow knowledge without belief, as in the case of the timid student who knows the answer but has no confidence that he has it right, and so does not believe what he knows.' Lewis (1996, p. 556)

[32] Or at least, we shouldn't accept knowledge without the epistemic state that epistemologists call 'belief'. This may be rather different from the ordinary English notion. So I do not take myself here to be in disagreement with e.g. Myers-Schulz and Schwitzgebel (2013). See Ch. 7.

> S knows that p just in case, for some known propositions $\{e_1, \ldots, e_n\}$, (i) S believes that p on the basis of $\{e_1, \ldots, e_n\}$, and (ii) all the cases where $\{e_1, \ldots, e_n\}$ are true are p cases.

On this approach, my knowledge depends on whether that which I may legitimately rely upon is strong enough to eliminate all the relevant alternatives; the conversational context of the knowledge ascription affects not only which alternatives are relevant, but also that which I may, for such purposes, be said truly to legitimately rely upon.

One might persist in worrying about circularity, given that 'known' appears in what is intended to be an explication of the truth conditions for knowledge ascriptions. But this worry would be misplaced. Note that I am not intending to offer an *analysis* of either knowledge or the concept KNOWS, breaking it into components that are conceptually or metaphysically prior to the analysandum; as I explained in the Introduction, my interests are much more in illuminating metaphysical roles that knowledge plays. Furthermore, the introduction of a basing requirement allows the treatment to fit much better with the knowledge-first approach to evidence. Lewis's account, with contextualist E=K, was all but vacuous; mine changes the shape of the view. Notice, for instance, that the statement above is not vacuous in the way the previous one was; it places a genuine constraint on knowledge: in order to know that p, you must believe that p on the basis of some sufficient knowledge. The framework fits naturally with a *recursive* approach to knowledge—there are two ways to be knowledge: a belief can be appropriately based in other knowledge, or it can be based on itself; this would be *basic* knowledge.[33] In other words, the approach invites a kind of *foundationalism*. Any kind of foundationalism invites two central questions: (1) what makes a given belief foundational, and (2) how do non-basic beliefs derive epistemic support from other supported beliefs? The Lewisian approach here has a straightforward answer to that latter question: the epistemic support in question comes from local entailment, i.e., entailment relative to a contextually-influenced set of relevant possibilities. But it is, for all we have said so far, wholly neutral on the substantive question of when it is possible to have basic, i.e., self-evident, knowledge.

The kind of contextualism I am advocating may have a tendency to obscure the importance of this question. Since in different conversational contexts, different epistemic standards will be in play, one might be tempted to suppose there's no objective foundation to belief; the only question concerns what is taken for granted in a given conversational setting. But this would be a mistake; contextualism is no substitute for substantive epistemological theorising. As I highlighted in the Introduction, contextualism is sometimes criticised for being too concessive to skepticism; in his provocatively-titled 'The Contextualist Evasion of Epistemology', Kornblith (2000) accuses contextualism of being irrelevant for responding to skepticism. While he admits that Keith DeRose, his primary contextualist target, does have a substantive response to skepticism, he observes that it derives more from a kind of substantive epistemic externalism than from contextualism. There is a limited sense in which I think Kornblith is right about this—there are many substantive epistemological questions to which contextualism alone is not the answer.[34] Contextualism concerns the

[33] An alternate formulation would have it that basic knowledge need not be based on anything at all:

> S knows that p just in case, either (a) S's belief that p is basic knowledge; or (b) for some known propositions $e_1, \ldots, e_n$, (i) S believes that p on the basis of $e_1, \ldots, e_n$, and (ii) all the cases where $e_1, \ldots, e_n$ are true are p cases.

I have elected for the simpler formulation given in the main text, as I prefer to think the case of basic knowledge as 'self-evident', rather than something that stands in no need of evidence. But I am not sure whether much hangs on this difference in possible descriptions.

[34] But as indicated in the discussion of Ernest Sosa's similar critique in the Introduction, I do not agree with Kornblith that this fact removes the interest—or even the *epistemic interest*—of contextualism; if contextualism is true, epistemologists must exhibit sensitivity to that fact, lest they equivocate. It's also rather plausible that there are additional respects in which

word 'knows'; much epistemology—to understate things—is not about this word. (Indeed, some epistemology does not obviously concern even knowledge itself—though if the knowledge first programme is correct, more epistemology depends on knowledge than is sometimes appreciated.)

Contextualist epistemologists, as much as any other foundationalist epistemologists, need to have something to say about basic knowledge: some beliefs can be held foundationally; which ones are they? Contextualists do not avoid this question, even though they might fail to notice that it is there. Lewis himself, for instance, committed rather quietly—without so much as obviously realising that it was a commitment—to a certain traditional internalist conception of basic knowledge. In the next two sections, I'll articulate this unrecognised commitment and express some reason to be uneasy about it, before going on to articulate an underappreciated alternative.

3.11 Lewis and Cartesian Contextualism

Here again is Lewis's characterization of evidence:

> [T]he uneliminated possibilities are those in which the subject's entire perceptual experience and memory are just as they actually are. ... [A] possibility W is uneliminated iff the subject's perceptual experience and memory in W exactly match his perceptual experience and memory in actuality. (If you want to include other alleged forms of basic evidence, such as the evidence of our extrasensory faculties, or an innate disposition to believe in God, be my guest. If they exist, they should be included. If not, no harm done if we have included them conditionally.) Lewis (1996, p. 553)

Although he lets it pass without remark, this amounts to a substantive epistemological commitment. In the framework of the previous section, Lewis is signing up for a phenomenal, internalist conception of evidence; this implies a parallel approach to basic knowledge. One way to see this is that on Lewis's view, there is a sense in which there can be no serious epistemological question about a subject's access to her own sensory experiences, the way there can always be a serious question about her access to external facts, even those in her immediate environment. For example, suppose that Angelina is in court, and she sees Edwin across the courtroom. Does she know that Edwin is present? On Lewis's view, to answer the question, we consider whether there are possibilities consistent with her visual experiences in which Edwin is not there; depending on whether we're countenancing such far-fetched possibilities, we may answer 'yes' or 'no'. But no such dependence on the context is relevant for considering whether she knows whether *she has the perceptual appearance* as of Edwin; it is built into Lewis's treatment of evident that her perceptual experience rules out *absolutely any* possibility in which she fails to have that experience; no matter which skeptical possibilities are relevant in a context, Angelina's evidence will always rule out all cases in which she doesn't have those experiences, and so she can satisfy 'knows that she has the perceptual experience as of Edwin' automatically. This is clearest on Lewis's simpler formulation, which omits a basing and a belief requirement—there, Angelina automatically satisfies the ascription no matter what; on my more complex requirement it may be possible for her to fail to satisfy this knowledge ascription in some contexts, for example, because she may not believe it, or she may believe it for the wrong reason. But there is still a sense in which there is no need to explain her epistemic access to the

a clear understanding of the contextualist semantics for 'knows' can impact our understanding of the theoretical significance for knowledge. Indeed, this is one of the central ideas of this book.)

fact—we might encode this by saying that she is always *in a position* to know all the facts about her sensory experience.[35]

Lewis's internalist treatment of evidence is not mandatory; one can insert one's own favoured foundational epistemology here. It is surprising that Lewis doesn't remark on making a commitment about a central choice point in epistemology. His remarks to the contrary notwithstanding,[36] I can only assume that he was motivated by something like the Cartesian idea that it is our own internal minds that we know first and best. Certainly the standard skeptical scenarios we trade in tend to be ones where subjects have the same subjective experiences they would in cases of genuine knowledge. Nevertheless, there are reasons to proceed carefully here.

For one thing, the motivation for the Cartesian idea that there is a kind of special security in our knowledge of the internal is somewhat dubious. Certainly, I think, the idea that there is some kind of special infallible privileged access to the internal has been decisively put to rest.[37] As Descartes famously observed in his *Meditations*, there are skeptical scenarios in which one is wrong about one's beliefs about the external environment, but where one's judgments about one's own experience are still correct. But if we admit what seems clear—that it is at least possible for subjects to be wrong about their own experiences—then there are also skeptical scenarios targeting the latter, at least in the sense that there are possible cases where one is wrong about them. So the characterisation of skeptical scenarios as cases where one's belief is false is inadequate. What must be added? The Cartesian idea is that a skeptical scenario is one in which one's beliefs are false, *and things seem exactly as they actually do*. But this is in effect to encode the assumption that seemings play a privileged epistemic role into the nature of what it is to be a skeptical scenario.[38] A more neutral characterisation of skeptical scenarios might have it that a skeptical scenario with respect to p is a possible case where one is wrong about p, even though one has all the same basic evidence. If basic evidence amounts to all and only facts about how things seem, then the Cartesian idea will be vindicated; but more externalist interpretations will also be available. So I don't think that reflection on Cartesian cases themselves gives *motivation* for the Cartesian internalist picture. I'll explore an externalist alternative in §3.12 below.

A more direct challenge to the Cartesian program Lewis signs up to is that it simply seems like a mistake to suppose there's no substantive epistemic work to be done with respect to knowledge of the internal. Eric Schwitzgebel's recent book *Perplexities of Consciousness* makes a strong case for a certain degree of skepticism about introspection; his case studies of first-personal knowledge concerning dreaming, the geometrical character of visual experiences, peripheral tactile experience, and others all add up to a case that 'people in general know very little about what might seem to

[35]There might be reason to complicate this even further, depending on just what we think it takes to be 'in a position' to know something. One might not wish to apply that description to some possible person who lacks the cognitive resources to base beliefs about her experiences on the facts about her experiences. For the distinction, see Ichikawa and Jarvis (2013, pp. 32–33).

[36]Lewis seems to think his phenomenal conception of evidence is a way of avoiding these difficult epistemic issues. He writes that if we focused instead on the *contents* of experiences, rather than the fact of their occurrence, 'we would need to tell some fishy story of how the experience has some sort of infallible, ineffable, purely phenomenological propositional content … Who needs that?' Lewis (1996, p. 553).

[37]For example, Timothy Williamson's famous 'anti-luminosity' argument puts serious pressure against the idea that any non-trivial state, including internal ones, are such that any subject in that state is in a position to know that she's in that state. Williamson (2000, Ch. 4). I do note however, that depending on just what is meant by 'infallible' and 'privileged', one might be able to maintain the official letter of the claim that there is infallible privileged access to the mental, if one embraces a kind of disjunctivism about introspection—see e.g. Macpherson (2010). This view admits that it's possible to make false judgments about one's internal experience, but denies that these share a common core with genuine introspective knowledge; so perhaps the latter count as in some sense infallible, since the grounds of such judgments entail their truth. Note however that disjunctivist views of perception (e.g. that of McDowell (1982) and Pritchard (2012)) have the corresponding result that perception is infallible, and it's not clear that being a disjunctivist about introspection but not about perception is well-motivated. So even disjunctivists should admit that there isn't *special* infallible privileged access to the internal.

[38]Neta (2003, p. 5) notices this assumption in traditional, Cartesian, forms of skepticism.

be obvious features of their stream of conscious experience'. Schwitzgebel (2011, p. ix) A detailed examination of these case studies is beyond my scope here, but if anything like Schwitzgebel's view even *could* be right, the Cartesian assumption Lewis takes on is highly questionable.[39]

None of this comprises anything like a definitive argument against the Cartesian approach. But I do think it highlights the respects in which it represents a substantive choice, and perhaps a less-than-fully-motivated one. To complete this case, it is helpful to sketch an alternative, externalist view. While I'm not sure which if either of these views is ultimately correct, I think that the possibility of the externalist interpretation has been unjustly ignored by contextualists to date. So I will concentrate on defending a kind of Moorean alternative interpretation of Lewisian contextualism. I turn to this task now.

3.12 Moorean Contextualism

The Moorean response to skeptical arguments is to claim knowledge of the negation of skeptical scenarios on the basis of knowledge of some ordinary, intuitively known, proposition. For example, if a skeptic offers this classic argument:

> (M–1) Edwin doesn't know he's not a brain in a vat
>
> (M–2) If Edwin doesn't know he's not a brain in a vat, he doesn't know he has hands;
> so
>
> (M–3) Edwin doesn't know he has hands,

the Moorean response is a straightforward instance of *modus tollens*:

> (M–4) Edwin knows he has hands
>
> (M–5) If Edwin doesn't know he's not a brain in a vat, he doesn't know he has hands;
> so
>
> (M–6) Edwin knows he's not a brain in a vat.[40]

'The Moorean response' itself is, as the term is typically used, this flat-footed response to skepticism; contemporary neo-Mooreans like Ernest Sosa and Duncan Pritchard supplement the Moorean response with a substantive epistemic story about how one manages to know the ordinary fact.[41] For example, a disjunctivist approach to perception will have it that the proper exercise of the faculty of perception directly delivers knowledge that one has hands.

I have used the canonical example of perceptual knowledge of hands in the face of the skeptical possibility that one is a brain in a vat; but there is a general argument *form* that is common to various kinds of skeptical puzzles. Compare for example, the lottery paradox and its Moorean response:

[39] On Lewis's own simple formulation, it is trivial that there just *is* no possibility that a subject might not know about features of his own conscious experience, for his having the experience trivially rules out every possibility in which he doesn't have it, and Lewis thinks this is sufficient for knowledge. On my more complex modification of Lewis, things aren't quite so straightforward—one *can* fail to know such facts, if one fails to believe them on their own bases, but it still seems rather implausible that such very difficult questions could comprise basic knowledge.

[40] The Moorean argument is, as the name suggests, associated with G. E. Moore and his famous 'Proof of an External World'. (Moore, 1939) The exegetical issues here are complex, however; most directly, Moore is engaging with the idealist, not the skeptic. Contemporary non-skeptical epistemologists have recognised that something like Moore's argument may be at least as attractive in engagement with the skeptic—and the resultant Moorean argument does fit reasonably well with what Moore *did* write about skepticism (e.g. Moore (1959)). For my current epistemological purposes, we needn't delve into the relationship between the historical G. E. Moore and 'the Moorean response to skepticism'; certainly contemporary epistemological literature has cottoned on to the interesting argument given in the main text above.

[41] I take the distinction and terminology from Pritchard (2012, p. 116).

(L–1) You don't know your lottery ticket won't win

(L–2) If you don't know your lottery ticket won't win, you don't know you won't be a millionaire next year; so

(L–3) You don't know you won't be a millionaire next year

versus

(L–4) You know you won't be a millionaire next year

(L–5) If you don't know your lottery ticket won't win, you don't know you won't be a millionaire next year; so

(L–6) You know your lottery ticket won't win.

Moorean and contextualist responses to skepticism are often described as if they are in conflict.[42] But as we have seen, this is a mistake. Contextualism is orthogonal to foundational questions about the basis for knowledge. One way to see this is to observe that the kind of Moorean approach gestured at here is an alternative to the Cartesian assumption *within the contextualist framework* discussed in the previous section.

Recall that contextualists, as much as anybody else, need to tell some kind of story about basic knowledge. Once we recognise that it is a substantive and non-mandatory choice to afford propositions about one's own sensory experience this privileged role, we can see new space for unexplored versions of contextualism. Consider, for example, the kind of disjunctivist approach to perception defended by John McDowell and others.[43] What is characteristic of this sort of approach is that cases of (genuine, veridical) perception yield direct knowledge of the external world. For example, when I raise my hands in front of my face and point my eyes at them, I *see that I have hands*; it is this world-involving factive mental state, as opposed to any purely intrinsic state, that provides my evidence for the belief that I have hands. This kind of disjunctivist approach is strongly associated with a kind of neo-Mooreanism in the sense articulated above.[44] The degree to which such a disjunctivist approach is well-motivated is well beyond my present scope, although I am of the opinion that it has considerable merits. The point I wish to make now is that such a Mooreanism is in no tension whatsoever with contextualism; nor need they be construed as competing explanations for the same phenomenon. On the contrary, a contextualist implementation of this Moorean idea can be well-motivated as an alternative to the Cartesian approach Lewis defended.

On this kind of 'Moorean contextualism', we will allow for basic knowledge of the external world via factive perceptual states like *seeing that one has hands*. In the framework I have been using, this amounts to allowing that some propositions—those to which one stands in factive perceptual states—can be known on the basis of themselves. This requires no departure from the letter of the Lewisian contextualism we've been working with; it simply invokes a richer class of basic knowledge/evidence.

The resultant view will have some features that are very different from those of Lewis's form of contextualism. The richer evidential base will have more anti-skeptical consequences. For example, suppose that Edwin sees that Angelina is in the courtroom. Consider this sentence:

(K_c) Edwin knows that Angelina is in the courtroom.

[42] E.g. in Pritchard's (2005) 'Neo-Mooreanism Versus Contextualism'.

[43] McDowell (1982) is an important articulation and defence of this approach; a more contemporary treatment is Pritchard (2012). There are helpful overviews and contemporary discussions in Byrne and Logue (2009) and Haddock and Macpherson (2008).

[44] Pritchard (2012, p. 116) is explicit in linking disjunctivism with neo-Mooreanism.

Now consider a very skeptical conversation context, where speakers are letting their epistemological fantasies rip—they take seriously the possibility that Edwin is a disembodied brain in a vat, or the possibility that Angelina has a hitherto-unknown identical twin who might be taking her place in court, or what have you. Lewis's version of contextualism will predict that in such a conversational context, (K_c) will express a falsehood. For Edwin's basic knowledge—the relevant facts about his sensory experience—do not rule out certain relevant possibilities in which Angelina is not in the courtroom. Fig. 3.2 illustrates Edwin's epistemic situation. The grey area represents the space of possibilities that Edwin's phenomenal evidence eliminates; white possibilities remain open. The circle represents the set of relevant possibilities—since we're discussing a skeptical context, it is a rather large circle, including the possibility that Edwin is a brain in a vat. The point marked '@' represents Edwin's actual situation.

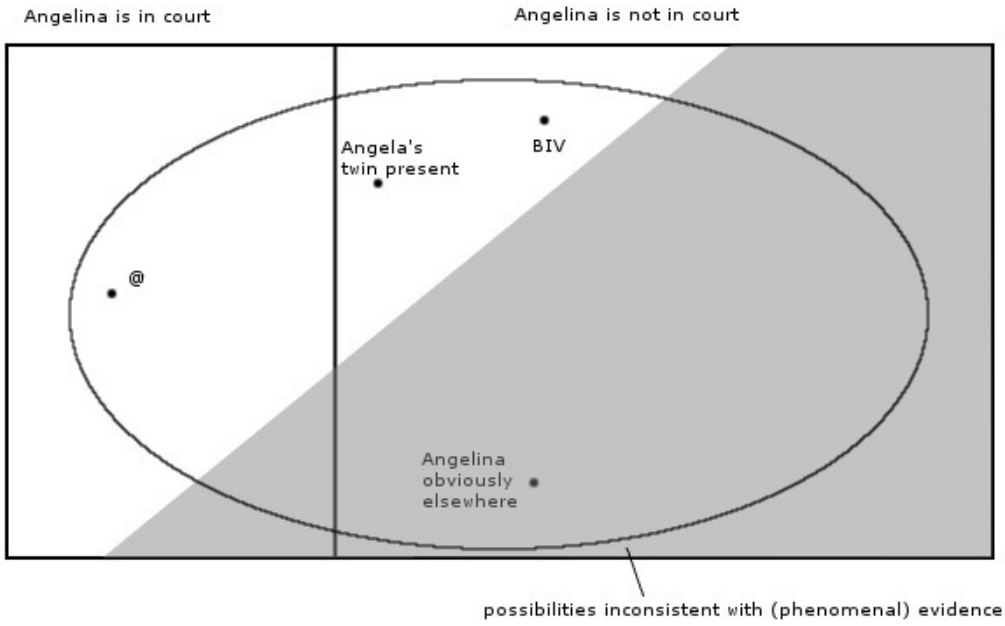

Figure 3.1: On a Cartesian evidential framework, there are uneliminated possibilities where Angelina is not in court. If they are relevant, as given by this rather skeptical set of possibilities, (K_c) is false.

On the Moorean form of contextualism, this skeptic-friendly result will not obtain. Sentence (K_c) expresses a truth, because Edwin's basic knowledge includes the fact that Angelina is in the courtroom, which rules out every relevant possibility in which she's elsewhere.

Although Moorean contextualism has this antiskeptical result, this is *not* because it holds that the pragmatics of conversation do not allow these skeptical possibilities into relevance. I am here proposing *no* changes to Lewis's own story about which possibilities are relevant, given a conversational context and the subject's situation.[45] Let's grant that the possibility that Edwin is a brain in a

[45]Though see §1.7, where I argue for unrelated reasons that we should modify his story in certain ways. The present point holds whether or not one goes along with me in these modifications.

vat *is* relevant—it isn't 'properly ignored'. On Moorean contextualism, that possibility is *ruled out* by Edwin's perceptual knowledge.

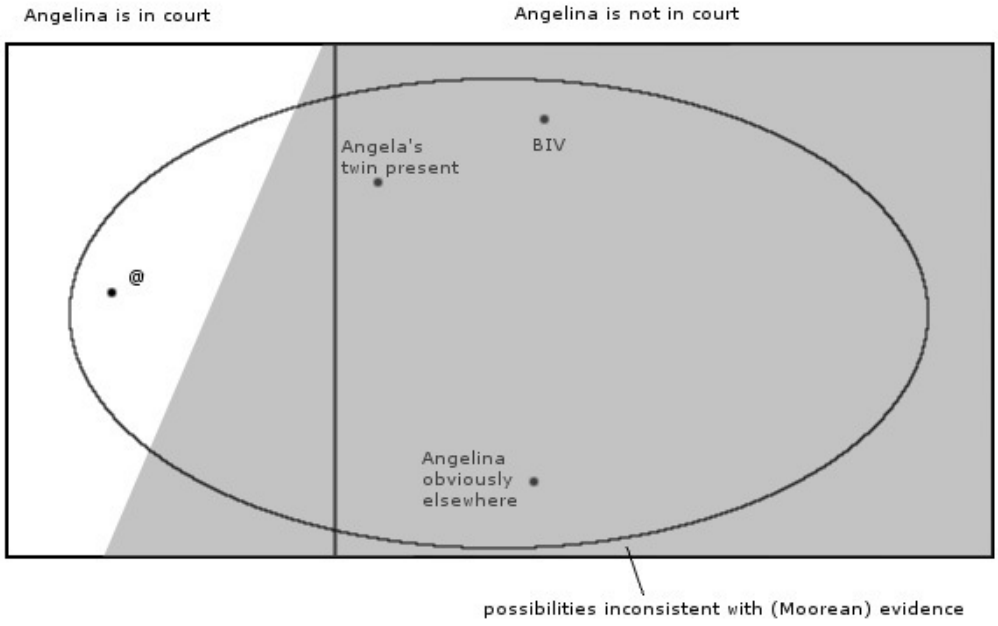

Figure 3.2: On a Moorean evidential framework, there are no uneliminated possibilities where Angelina is not in court—not even if radical skeptical scenarios are relevant. For Edwin's evidence includes the seen proposition that she is present.

Indeed, Moorean contextualism, unlike Lewis's Cartesian contextualism, predicts the truth of (K_b), relative even to the skeptical contexts where it might be uttered:

(K_b) Edwin knows that he's not a brain in a vat.

3.13 Skeptical Intuitions and Moorean Contextualism

The last section showed that Moorean contextualism straightforwardly mimics the traditional Moorean response to radical skeptical worries. But Moorean contextualism also departs from traditional Mooreanism in important respects. Consider for example the lottery paradox. Here again is its Moorean treatment:

(L–4) You know you won't be a millionaire next year

(L–5) If you don't know your lottery ticket won't win, you don't know you won't be a millionaire next year; so

(L–6) You know your lottery ticket won't win.

The first premise here is that *you know you won't be a millionaire next year*. This premise is in some respects very *unlike* the key anti-skeptical premise in the BIV argument, that *Edwin knows*

he has hands. The latter is, on plausible disjunctivist stories, delivered directly by perception; the former isn't. Even if you do know that you won't be a millionaire next year, this is *not* because you're acquainted directly with the fact by seeing it. One can't *see* how much money one will have next year in the way that one can see how many hands one has.[46] So facts about how much money one will have next year are not candidates for basic knowledge.

What this means is that the Moorean contextualist will *not* mimic the traditional Moorean with respect to the lottery paradox, the way she will with respect to the BIV paradox. Relative to conversational contexts in which the possibility that your lottery ticket will win is relevant, (L–4) and (L–6) are both false, since none of your evidence rules out certain relevant possibilities where your ticket wins and you become a millionaire. But if you see that you have a ticket, then there may be no contexts, no matter how skeptical, in which you're addressed but 'you know that you have a lottery ticket' fails to be true.[47]

Borrowing some terminology from Stewart Shapiro, we may say that basic knowledge is a kind of 'super-knowledge'. Shapiro explains the relevant schema thus:

> An argument form is super-valid just in case it is valid in all legitimate structures and theories. Notice that a similar stipulation can be made for most context-sensitive terms, no matter how the context-sensitivity is articulated—contextualist, assessment-sensitive, whatever. A food is "super-tasty" if it is tasty to everyone; an object is "super-left" if it is on the left according to every vantage point; a person is "super-ready" if she is ready for *anything*; a person is a super-enemy just in case he is an enemy of everyone; a pub is super-local if it is local to everywhere. In many cases, of course, the coined term will be completely useless. A super-local pub would be quite handy, but, alas, there aren't any. (Shapiro, 2014, p. 114)

Basic knowledge is that which is super-known. On Lewis's Cartesian picture of evidence, I super-know about my phenomenal states. Even relative to the most skeptical contexts, 'I know that I am undergoing an experience as of a green light' is true. On the Moorean picture of basic evidence, I super-know that I have hands; because my basic evidence entails that I have hands, there is no context so skeptical that it introduces uneliminable possibilities in which I have no hands. But anything that goes beyond that which I can see for myself—my knowledge that I will have hands tomorrow, for example—cannot be super-known. Even if 'I know I will have hands tomorrow' is true according to most ordinary contexts, there are some skeptical contexts that countenance possibilities in which I lose my hands overnight; since these are cases my evidence can't rule out, I don't know that they won't obtain.

It is not *mandatory* for a Moorean contextualist to go this way; one *could* allow that one can even have basic knowledge of contingent matters about the future. At the extreme, one might even hold that everything satisfying 'knows' in *any* context is basic. If one did that, the contextualism would be idle, as one's basic evidence would entail the beliefs in question relative to every context. An underexplored and cryptic remark by Timothy Williamson in *Knowledge and Its Limits* is suggestive of this kind of approach: 'E=K suggests a very modest kind of foundationalism, on which all one's knowledge serves as the foundation for all one's justified beliefs.' (Williamson, 2000, p. 186). I am not sure how seriously Williamson intends this remark; read literally, it suggests that all knowledge

[46] If there are facts about the future and crystal balls give transparent access to them, perhaps one can see what will happen next year. We may stipulate, however, that you have no such crystal ball.

[47] An alternative approach, similar in spirit to my own, is that given by Ram Neta in his (2003). Neta's approach has it that what shifts according to context is that which is counted as evidence; in nonskeptical contexts, the external-world deliverances of perception are counted as evidence, but in skeptical contexts, our ideas about evidence are more restrictive—perhaps, for instance, we consider only propositions about how things seem. On my view, even if Neta is right that we tend, in skeptical contexts, to think our evidence is so restricted, we do so in error.

is basic knowledge—none of it depends for its justification on any other justified beliefs.[48] The idea that some knowledge is more basic than other knowledge, however, is an extremely intuitive one. My knowledge that I have hands right now as I type is a much better candidate for being basic than is my knowledge that I will have hands in a few months when I submit this manuscript. And I think this intuitive difference is reflected to some degree in the robustness and naturalness of skeptical intuitions. I shall return briefly to the possibility that all knowledge is basic in §4.9.

When I introduce skepticism to the students in my introductory epistemology courses, it takes a bit—not a lot, but a bit—of work to get them to feel the force of the Cartesian evil demon. I have to emphasise to them that we're describing a possible case where *things seem exactly the same*, and that there's no obvious grounds on which we can rule it out. After a few minutes along these lines, most of my students can feel the skeptical pressure, although I have the sense that they tend to think of it more as a philosophical game than as a serious threat to the reasonability of their perceptual beliefs. My conversations with others who teach this material suggest to me that this is not an atypical reaction.

By contrast, my students come in on the first day of class willing to assert unprompted, as if it were obvious, that one can never know what's going to happen in the future. Here, the work I need to do to help them appreciate the puzzle is to tease out the *anti*-skeptical commitments they already have. (When I ask them whether, having read they syllabus, they think they know whether I'll be giving a final exam, they admit that it seems they do.) The idea that you can't know what you can't see for yourself is a very intuitive one; the idea that you can't know what you can't prove in a way that would satisfy the Cartesian skeptic takes a bit more effort to motivate. If you doubt this comparative claim, I suggest an experiment: go outside and find some passerby willing to spend a minute philosophising with you, and play the skeptic. Hold up a copy of *Harry Potter and the Chamber of Secrets* and try to convince some of them that they don't know who wrote it. ('J. K. Rowling might have stolen credit from someone else.') And try to convince others that they don't know whether there are any books nearby. ('This might be some other object that just looks a lot like a book.') I predict that, assuming you find enough patient individuals, you'll have much greater success in the former case.

Note also that the Moorean way with the skeptic is much more plausible in the case of plausibly basic knowledge than it is with inferential or derivative knowledge. While it's notoriously unsatisfying to rest one's knowledge that one isn't a brain in a vat on one's knowledge that one has hands, it's even *less* satisfying to claim knowledge that one's lottery ticket will lose on the grounds that one won't be able to afford to go on safari next year.[49] Similarly, I cannot know that I won't have a fatal accident this year on the grounds that I will teach philosophy of religion next spring. (Maybe I know it, but not on these grounds.) Some instances of the Moorean schema are very implausible. So a Moorean should restrict that response to the skeptic to plausible cases of basic knowledge. These include the deliverances of perception, but exclude contingent facts about the future.

The Moorean contextualist who distinguishes basic from nonbasic knowledge mimics the traditional Moorean with respect to skeptical worries that challenge the direct results of perception, but mimics the Cartesian contextualist with respect to challenges to these latter cases of apparent ordinary knowledge. Relative to skeptical contexts where the possibility that someone other than

[48] I take this to be the standard definition of foundationalism—it is the claim that there are justified beliefs that do not themselves depend on justification for any other belief. A knowledge-first proponent will emphasise the theoretical importance of knowledge over justification, but this is still consistent with the idea that the justification (indeed, the knowledgeability) of some beliefs depends on that of others. Read literally, this passage seems to suggest that all knowledge is epistemically on a par. This strong claim is not emphasised elsewhere in Williamson's work. But see §4.9.

[49] The example is from Hawthorne (2004, pp. 1–3). Note that this is more a point about warrant transmission than one about closure. See Wright (2002). It *is* intuitive that if one knows one won't be able to afford to go on safari next year, then one knows that one's lottery ticket will lose. What is intuitively very *im*plausible is that one can know that one's ticket will lose *on the basis of* the claim about the safari. Perhaps one knows that one's ticket is a loser, but *this* couldn't be the reason.

Rowling secretly wrote the *Harry Potter* books is relevant, 'S knows that Rowling wrote the *Harry Potter* books' would require S to have special evidence that most people do not have. For example, if S herself *saw* Rowling writing those books, then she would know even by the high standards.

Notice that something like the epistemic primacy of seeing something for oneself is present in some of Keith DeRose's discussions of more skeptical epistemic standards. He writes, of one of his cases,

> Thus even though Thelma has grounds that would usually suffice for claiming to 'know' that John was at the office yesterday (she heard from a reliable source that he was in, though she doesn't know what exactly that report was based on, and she herself saw John's hat in the hall), she is in her present context admitting that she does 'not know' that he was in—though, she adds, she has good reason to think he was in. In this context, she would claim to 'know' he was in only if, say, she herself has a clear recollection of having herself seen him in. DeRose (2005, pp. 186–7)

The idea that one has privileged epistemic access to that which one has perceived oneself is a very natural one[50]; I see no reason to prefer the Cartesian view that one has privileged access only to one's own mental experiences. Introspection is often much more difficult than ordinary perception is. There are admittedly important challenges and questions for the disjunctivist approach I have been describing. For example, just which kinds of things can we directly perceive? How theory-laden is perception? I have been writing as if one can directly *see that one has hands*; should one instead say that strictly speaking one can only directly *see that there are hand-shaped objects*? The latter approach would still give direct access to the external world, but only to weaker properties of it. Is there non-perceptual basic knowledge? These are all important questions. But notice that these questions are not obviously more difficult than the parallel questions that often go unnoticed with respect to more traditional Cartesian approaches to perception: to just which aspects of one's mental experience can one have direct access? Again, everyone has to say something about basic knowledge, so everyone faces questions of this kind.

I cannot hope in this volume to give anything like a full explication and defence of the disjunctivist epistemology I have been sketching. But I do hope that I have said enough to make it clear that it represents a possible approach to basic evidence. While I'm not *convinced* that the Moorean version of contextualism is preferable to the Cartesian one, I am attracted enough to it to sign up for it as the official view of this book. Where it becomes relevant in future chapters, I'll mention that my approach could be developed in a different, more Cartesian, way.

3.14 Radical Skepticism

Signing up to this form of Moorean contextualism is a further development—not a modification—of the contextualist line developed earlier in the book. None of the cases I have relied on in motivating contextualism have essentially traded on skeptical scenarios that basic perceptual knowledge rules out. For reasons like those articulated in the previous section, I do not consider such *radical* skepticism to provide the strongest challenges to knowledge (in this, I disagree with Kornblith (2000)[51]); so I do not think providing a resolution of them is a central motivation for contextualism. But I should pause at least briefly to discuss radical skeptical scenarios.

[50] Turri (2014a) gives empirical support for the idea that among the folk, direct perceptual knowledge is perceived as less vulnerable to skeptical challenge than inferential knowledge is.

[51] See also Brown (2006, p. 411) for similar commitments to the idea that contextualism is fundamentally tied up with radical skepticism.

As indicated in the previous section, my view has it that any scenario in which the contents of my actual direct perception are false, even ones in which things seem internally just as they actually are, is ruled out by my evidence. So, insofar as it is intuitive that one can't know that one isn't a brain in a vat, my brand of contextualism has no direct way to accommodate that intuition. Now as I've been indicating, I think it is an advantage, not a cost, of this approach that it treats such *radical* skeptical worries as importantly different from skeptical worries in more mundane contexts—I am not sure the skeptical verdicts about these cases really *are* as intuitive as Lewis assumes they are[52]—nevertheless, it does seem reasonably clear that there is at least *some* possible pull in the direction of the idea that I can't know that I have hands, because I might be a brain in a vat. So again, insofar as these radical skeptical thoughts are intuitive, it would be helpful to have something to say by way of explaining those intuitions away.

Here, I think that the Moorean contextualist can helpfully take on many of the ideas that have been developed by the more traditional Mooreans in explaining away skeptical intuitions. For example, Jennifer Nagel has argued, in a series of recent publications, that thinking about possibilities for error often, as a matter of psychological fact, introduces doubt in subjects; this doubt is inconsistent with knowledge, not because the standards have been raised, but because the loss in confidence is inconsistent with outright belief.[53] There is no reason a Moorean contextualist couldn't run a similar line to explain some of these more radical skeptical intuitions—when confronted with the brain-in-a-vat hypothesis, one might undergo a kind of epistemic vertigo that induces failure to believe that one has hands. (So on this kind of view, perhaps epistemology *really does* destroy knowledge.[54]) The story will be less clear for third-person ascriptions; Nagel's move is to invoke a psychologically robust tendency to mistakenly project our own epistemic doubts onto others. On the Moorean line, another alternative is also available with respect to the objects of direct perception: one might, when thinking about skeptical scenarios, come to doubt the Moorean story itself. If I become agnostic about whether we're all brains in vats, I may become a Cartesian, rather than a Moorean, and mistakenly suppose that even in the good case, one doesn't have direct knowledge of the external world. So, while I am doubting, I may argue that whatever the subject's epistemic position, she doesn't know that she has hands because her evidence, which I'm taking to be exclusively internal, doesn't rule out the possibility that she's a brain in a vat. One advantage of this error theory over Nagel's is that it does not require me to make false judgments about the subject's own subjective level of confidence.

A natural challenge at this point is that, if I'm willing to take on invariantist strategies to explain away certain skeptical intuitions, why shouldn't I embrace a more traditional, full-blooded, invariantist version of Mooreanism? I admit that I am in a slightly precarious dialectical situation—I want to countenance a certain degree of Moorean externalism, but no more—nevertheless, I think that the position can be motivated. One motivation I've mentioned already: it does seem as if the intuitions I'm happy to explain away are much less clear than are more moderate skeptical intuitions, like the intuition that Carrie doesn't know whether the dog is at home (since after all, a burglar could have forced open the door and let the dog out in the half hour since she left her there).

One could also develop the Moorean line of thought in a skeptical invariantist direction: although I can directly see for myself that I have hands, I have no way to rule out the possibility that I'll lose my hands in a freak accident overnight. So I know that I now have hands, but I can't know whether I will have hands tomorrow—nor can I know anything that doesn't follow deductively from that

[52] I have been unable to find any experimental philosophy results that bear directly on this question. Turri (2014a) argues that perceptual knowledge is more resistant to moderate skeptical challenges than indirect knowledge is, but does not invoke radical skeptical scenarios. See also Dummett (1973, p. 355), who suggests that we learn the practice of assertion by focusing initially on facts we can recognise for ourselves.

[53] I'll investigate this kind of line in more detail in Ch. 7.

[54] See fn. 19.

which I can recognise directly. (I don't know whether my car is where I left it, or whether my dog has sprouted wings in the past hour, or where I was born, or whether Abraham Lincoln was an alien from Mars…) Given the neo-Moorean approach to basic evidence to which I'm attracted, this does seem to me to be the best form of skepticism available.[55] I do not embrace it because contextualism seems to me preferable: in many less skeptical contexts, it is intuitive that 'I know that Abraham Lincoln was human' expresses a truth.

Remember again that the precarious situation of adopting anti-skeptical strategies to some worries but not to others is ultimately one shared by *all* foundationalists, including the more familiar Cartesian ones. After all, it's not impossible to construct something that at least looks like a skeptical scenario with respect to introspective judgments too. For example, suppose that Edwin hasn't shaved since yesterday, and so has some stubble. Angelina looks at Edwin, and sees that he has stubble. Moreover, when she introspects, she notices that she has the perceptual experience as of seeing someone with stubble. One can challenge her perceptual knowledge by invoking a Cartesian skeptical scenario like this one:

> *Demon.* Edwin doesn't really exist; Angelina isn't really looking at anyone with stubble. Her experience as of seeing someone with stubble is caused by an evil demon who is directly manipulating her perceptual experiences.

This Cartesian scenario challenges her perceptual knowledge, but takes her introspective knowledge for granted. But one could challenge the latter with an even *more* radical skeptical scenario:

> *Superdemon.* Not only is Angelina not really looking at anyone with stubble; she doesn't really even have experiences *as if* she is looking at someone with stubble. Her introspective judgment of having such experiences is caused by an evil superdemon who is directly manipulating her introspective judgments.

The Superdemon scenario is a challenge to the nonskeptical verdict that Angelina knows that it appears to her as if Edwin has stubble. It is a possible case where it doesn't so appear, and it is one which, according to the challenge, Angelina's evidence is unable to rule out.[56] The Cartesian considers this challenge to be illegitimate, perhaps along lines such as these: the scenario in question, while possible, is one in which things seem different to Angelina; so her evidence—the way things seem to her—*does* rule out the Superdemon case. This seems like a reasonable response to the Superdemon case. But notice that it is exactly parallel to the Moorean response to the standard Demon case: the scenario in question, while possible, is one in which Angelina doesn't see the same things she actually sees; so her evidence—the contents of her factive perceptual states—*does* rule out the Demon case. So the Cartesian has to explain away certain skeptical scenarios too; he posits a difference between Demon and Superdemon, where I posit the relevant difference instead between Demon and the challenge to Angelina's inference that Edwin hasn't shaved today that we might call 'Tonic':

> Tonic. Edwin did shave today, but he used a special tonic that made him grow new stubble very quickly, in an attempt to deceive Angelina into believing he hadn't shaved today.

[55] For a defence of this sort of approach, see Davis (2007) and Davis (2017). Dodd (2007) argues that Williamson's E=K commits him to a view like this one.

[56] A reader for OUP helpfully suggests, citing Williams (2001b, ch. 8), an alternate way of setting out the challenge, taking as its starting point the thought that it's possible to be wrong about the contents of one's experiences—one might take oneself to have the experience as of p, but be mistaken: perhaps it's an experience of q instead. Since in general, one can be wrong about the contents of one's experience, it's possible to be misled by a superwomen.

Tonic challenges Angelina's inferential knowledge that Edwin didn't shave today. On my Moorean contextualism, Tonic represents a legitimate skeptical challenge in the way Demon and Superdemon do not. Assuming that Angelina has performed no checks that would exhibit sensitivity to the Faker hypothesis, 'Angelina knows that Edwin didn't shave today' is true only relative to contexts in which the Tonic hypothesis is properly ignored.

So everyone needs to draw a line somewhere. The Cartesian draws it between Demon and Superdemon; I draw it between Tonic and Demon. I see no reason to consider this Moorean stance ill-motivated.

Chapter 4

Justification

This chapter takes up *epistemic justification*. I'll articulate and defend a knowledge-first approach to justified belief, as well a contextualist approach to 'justified'. My approach to justification is similar to my approach to knowledge, but I want to begin by emphasising a difference between justification and knowledge. 'Knows' is an *extremely* common word—as Jennifer Nagel points out, is one of the ten most common verbs in the English language, and a rare *linguistic universal*, having a translation in every known language.[1] By contrast, 'justification' is to a considerable degree an epistemologist's term of art.[2] One quick way to see the difference is that, in an introductory epistemology course, one needn't do anything in particular to get students to understand what is meant by someone knowing something; but one must do some expository work to help them catch on to what we mean when we talk about epistemic justification. Some epistemologists, such as Stewart Cohen, are skeptical that philosophers have managed to latch on to a common subject matter in the literature on epistemic justification.[3] While I do not share this skepticism, I think that Cohen is right to emphasise that it would be a mistake to look to ordinary language usage to justify theories of justified belief. This marks a dramatic contrast with the case of 'knows'. Theorising about justification owes much less to ordinary intuition than does theorising about knowledge; it is more of a matter of *explication* in the Carnapian sense.[4] I'll defend a theory of justification and 'justified' that I take to best make sense of how epistemologists do *and should* think and speak.

An operating assumption of this book has been that the knowledge-first program is worth pursuing; the idea that knowledge has important theoretical roles to play in various domains is a fruitful and promising one. I do not assume the same about justification. I am much more agnostic about whether or to what degree epistemic justification will appear in our best theories of thought, action, etc. (I certainly do *not* expect justification to appear in our best theory of knowledge.) One might naturally wonder, then, how much value there is in articulating a theory of epistemic justification;

[1] Nagel (2014, p. 6), citing Wierzbicka (1996).

[2] I agree with Reynolds (1998, p. 533) on this matter; in fact, the view he develops is in some respects similar to mine. See §4.7. See also Wedgwood (2008, p. 7).

[3] Cohen argued for this conclusion at a workshop on intuitions in Oslo in 2009; he also defends it in Cohen (2016), available at http://www.stew-cohen.com/papers/as of July 3, 2015.

[4] Compare Edward Craig's remarks on the methodology of the theory of knowledge:

> We take some prima facie plausible hypothesis about what the concept of knowledge does for us, what its role in our life might be, and then ask what a concept having that role would be like, what conditions would govern its application ... then see to what extent it matches our everyday practice with the concept of knowledge as actually found. (Craig, 1990, pp. 2–3)

See also Haslanger (2000) and (2006). In fact, I think that *all* theorizing is in this sense explication to a degree; but this is particularly true in cases like that of justification, where the body of data in question is much thinner.

nevertheless, I do think there's reason to expect it to be a worthy project. For one thing, developing a theory of justification may help us to *discover* theoretical roles that it might be well-suited to play. For another, careful theorising about justification may well aid in the investigation of the knowledge-first program. Very often, for instance, critics of the latter will suggest that justification is better suited to play the role posited for knowledge.[5] Without an operating theory of justification, it is difficult to evaluate this kind of claim—especially since, as a theoretical term of art, there are many cases in which it is intuitively unclear whether it applies. And without knowing whether justification should be understood in terms of knowledge, it is unclear what bearing that claim would have, if true, on the knowledge first program. If justification itself is understood only in terms of knowledge, then if, say, reasonable action should be understood in terms of justified belief, it *follows* that it should ultimately be understood in terms of knowledge.[6] So a knowledge-first theory of justification would be a valuable thing. I have defended such a theory in my (2014) 'Justification is Potential Knowledge'. I'll repeat many of the arguments and ideas from that paper in this chapter, although my present view does differ from the view of that paper in key ways—in particular, I am now much less committed than I was then to a strong form of *internalism* about justification. (My change in this respect has much to do with the considerations discussed in §3.12.) I'll emphasise my own change in view in the relevant sections below. I also did not develop any kind of contextualism in my previous paper; I will do so now.[7]

4.1 Initial Clarifications

Because 'epistemic justification' is a philosopher's term of art, it is worth beginning with some gestures towards exemplars, in order to help fix the subject matter. When I introduce the idea of justification to my epistemology students, I start with some obvious cases: sometimes, there is overwhelming evidence in favour of some proposition, and so the appropriate thing to do is to believe it. For example, if you look outside and see that it is sunny, then you should believe that it is sunny. If you do believe that it is sunny on this basis, you've done a good job—your belief is *justified*. By contrast, some beliefs are unreasonable. If you just have a hunch that a fair coin is going to land tails, and believe according to that hunch, your belief is unjustified: that's not the kind of thing it's reasonable to believe. These are easy cases, but there are harder questions too. There's more controversy, for instance, whether any of *these* beliefs are justified, given the evidence available to us:

- The belief that there is a God who created the universe
- The belief that there is not a God who created the universe
- The belief that human activity is contributing to climate change
- The belief that eating meat is morally wrong
- The belief that there is a metaphysically significant distinction between past and future times

[5] See for example Comesaña and Kantin (2010) and Rizzieri (2009), cited and criticised in Littlejohn (2012, p. 106), for instances of this move as applied to the thesis that evidence is knowledge. Lackey (2008) and McKinnon (2015), apply the move to the knowledge norm of assertion.

[6] At an extreme, one might hold, as Sutton (2007) does, that knowledge itself is both necessary and sufficient for justification; if so, positing justification to play the roles offered to knowledge is no alternative at all. It is just along these lines that Littlejohn (2014) argues that a justification norm for action is consistent with a knowledge norm.

[7] This latter is *not* a change in view; I did always intend to go on to develop the approach in a contextualist way, but I found, when writing that paper, that the paper had enough commitments in the object language, so I restricted my focus accordingly. But the view of that paper is wholly consistent with contextualism about 'justified'.

The task of a theory of justification is to posit general principles that *unify* and *explain* why some beliefs are justified, and others aren't. A natural methodology for this project is to begin with the exemplars, and consider possible theories that explain what the clear cases of justified belief have in common, that may distinguish them from the clear cases of unjustified belief. Ultimately, one may hope from there to use the resultant theory to consider some of the more controversial cases.[8]

The kind of justification I am talking about is a property of *beliefs*. It is useful in this context to distinguish *doxastic* from *propositional* justification, where the latter concerns what belief (or other doxastic attitude) the totality of the subject's evidence supports, whether or not the subject forms such a belief.[9] It is very plausible that there are strong connections between propositional and doxastic justification, but I won't assume any particular connection at the outset. I won't assume, for example, that what it is to be doxastically justified is to base one's belief upon one's propositional justification.[10]

4.2 Desiderata for a Theory of Justification

It is useful to examine a few of the prominent past attempts to provide a theory of justification; observing their particular shortcomings will help us to develop a list of desiderata against which we might test my own preferred view.

On a certain form of *classical foundationalism*, justified beliefs are those which ultimately rest on internal states to which a subject has infallible access. The main problems with this traditional view are that (a) few if any internal states do seem to be such that one can infallibly recognise their presence or absence, and (b) few if any of our intuitively justified beliefs about the external world seem to be sufficiently well-supported by any decent candidates for such internal states. As a theory of justification, classical foundationalism is *too stingy*; it denies justificatory status to too many beliefs that are intuitively justified.

A less stingy approach in the neighbourhood of classical foundationalism would relax the requirement of infallible access. For example, Richard Fumerton (1995) develops a form of internalist foundationalism according to which, although foundationally justified belief does not require infallibility, it requires direct *acquaintance* with the fact believed. In many cases of intuitively justified belief, Fumerton suggests, such acquaintance obtains; for unjustified beliefs, there is no such acquaintance. Acquaintance is understood not to entail infallibility, so Fumerton's view is less stingy than the classical approach; this is an advantage. Nevertheless, many epistemologists, including myself, find the approach deeply unsatisfying, absent a further theory of what acquaintance consists in. Fumerton tells us only that he treats it as a *sui generis* relation, and offers some examples of its extension; this is not, to my mind, particularly clarifying of either acquaintance or justification. A theory of justification should be *illuminating*; it should improve our understanding of justification.

I tend to find a similar feeling of dissatisfaction with Earl Conee and Richard Feldman's (1985) 'evidentialism', according to which one is justified in believing that p just in case believing that p is the attitude towards p that best *fits the evidence*. Without a suitably independent grip on when and

[8] This isn't to say we're going to end up with a decision procedure anyone could follow to guarantee justified belief in all cases. Even armed with a correct theory of justification, we may or may not find ourselves with the resources to determine whether, for instance, atheism is a justified belief, given our evidential situation.

[9] In Ichikawa and Jarvis (2013), Benjamin Jarvis and I argue that both of these senses of justification should be distinguished from a third, which we call *ex ante* justification. This is roughly the property of being *in a position* to form a doxastically justified belief. That this state differs from propositional justification is one of the central (and controversial) ideas in that book.

[10] One reason I don't make this assumption is that I think it's false. For one thing, in some cases of *a priori* justification, propositional justification obtains automatically; in such cases, I think, there's nothing for a doxastically justified belief to be based *on*. Ichikawa and Jarvis (2016) argues that for reasons like these, propositional and doxastic justification need to be rather theoretically distinct.

why believing a content is the attitude that best fits the evidence, this is not an illuminating theory of justification.[11]

The desideratum that a theory of justification should be illuminating is consistent with the idea that such a theory will be a metaphysical theory about the nature of justification—it does not require any commitment to the idea that the project of a theory of justification is an 'analysis' of the concept JUSTIFICATION. A good metaphysical theory won't merely state a truth; it will improve our understanding. That's why e.g. the theory that *justification is that which obtains when a belief is justified* is a bad theory, even though it's plausibly true.

Epistemological *externalists* suggest that justification is not wholly a matter of a subject's internal states; they hold that whether a belief is justified might depend on features of the subject's external environment—perhaps even features of which she is unaware. For example, according to a simple form of reliabilism, a belief is justified if and only if it is produced by a reliable mechanism—a mechanism that tends to produce true beliefs rather than false ones.[12] One of the appealing features of reliabilism is that it unifies and explains what many of our intuitively justified beliefs have in common, and distinguishes them from many of our intuitively unjustified beliefs. And the notion of reliability appealed to, by contrast with Fumerton's *sui generis* 'acquaintance', seems to be much more clearly explicated and well-understood.[13] Furthermore, reliabilism fares much better than does classical foundationalism with respect to stinginess: it is plausible that many of our belief-forming mechanisms are actually reliable. So on reliabilism, it is plausible that many of our ordinary beliefs are justified, even if they are not derived from infallible beliefs.

Still, some critics have worried that reliabilism and other externalist views are still a bit stingier than they ought to be. They seem to imply, for instance, that a subject who is in some intuitive sense *doing everything right from her own point of view* might still fail to have a justified belief, due to unusual external circumstances' failure to cooperate. If, for example, Adam forms beliefs about the townspeople's comings and goings on the basis of apparently authoritative and reliable testimony in the tavern, it's reasonably intuitive that his beliefs will be justified. ('Zorah came back to town yesterday,' someone tells him, and so Adam believes that Zorah came back to town yesterday.) If it turns out that the appearance of authority and reliability is illusory—if it turns out, for instance, that it's a skilled con artist attempting to deceive Adam—then externalist views might have the result that this could make for a difference in justification. For example, a reliabilist might say that although Adam *thought* that reliance on this testimony was a reliable belief-forming method, it in fact isn't, and so in fact, his resultant beliefs are unjustified.[14] Insofar as it is intuitive that his beliefs *are* justified—just as justified as they would have been if the source of testimony were as honest as he seemed—reliabilism seems to give a counterintuitive result.

One strong form of this kind of internalist intuition is the 'New Evil Demon' intuition, due to Cohen (1984). If whether a belief is justified depends even in part on how things are outside the agent, then two intrinsic duplicates can differ with respect to the justification of their corresponding beliefs; but this is thought by internalists to be counterintuitive. For example, supposing that my perceptual

[11] Two notes: first, although I find their approach unilluminating, I do not claim that Conee & Feldman's view is platitudinous; it may well place some substantive constraints on justification, such as the supervenience of justification on evidence. I don't think that's enough, although I do think it's something. Second, Conee & Feldman are to a first instance developing an approach to *propositional*, rather than *doxastic*, justification. But they do give a more complicated theory of the latter in terms of the former; on that score, their view is a competitor to the one I'll be developing.

[12] The locus classicus is Goldman (1979).

[13] Or at least, it enjoys an advantage on this score if there's a suitable answer to the 'generality problem'. See Conee and Feldman (1998).

[14] It is a complicated matter whether a reliabilist *would* say this; it depends on how coarsely one individuates methods. (Perhaps the relevant method Adam is using is reliance on testimony, rather than reliance on this particular individual's testimony; maybe, then, it is still on the whole reliable enough for justification.) This question again is an instance of the generality problem for reliabilism.

beliefs are actually reliable, reliabilism has it that my perceptual beliefs—e.g. my present belief that it is sunny, based on my perceptual experiences—are justified. But my twin, the unfortunate victim of an evil demon, made to have misleading experiences as of sunniness, is not justified in believing that it is sunny. Although externalists have offered many attempts to soften the blow of denying the New Evil Demon intuition,[15] having to do so does seem to remain a cost; a theory of justification will have a mark in its favour if it can respect these internalist intuitions.

The New Evil Demon intuition is a strong generalisation on the kind of internalist intuition invoked about Adam's testimonial belief mentioned above; the latter held that in a particular sort of case, particular kinds of differences in the external environment, without a corresponding change in the subject's subjective experience, can't make the difference for whether a subject's belief is justified; the former has it that *in general*, differences in the external environment without a corresponding change in the subject's subjective experience can *never* make the difference for whether a subject's belief is justified. In my previous presentation of these ideas in Ichikawa (2014), I was moved to try to accommodate even this strong form of the internalist intuition. I am now less concerned with this general claim, although I continue to think it important to respect many instances of the internalist intuition, like that of Adam's. The reason I'm now less moved by the more theoretical generalisation has to do with my reflections in Ch. 3 on factive perceptual states and basic knowledge. I'm now attracted to a neo-Moorean disjuntivist treatment of perception, according to which in favourable circumstances subjects have direct perceptual epistemic access to features of the external world. On this picture, when I look outside and *see that it is sunny*, I have direct perceptual access to facts about the weather. While this view is *consistent* with the New Evil Demon intuition, a disjunctivist might well wonder what motivation there is to treat that intuition as true. My twin who is deceived into believing that it is sunny is, on this view, unlike me in key epistemic respects. In particular, he lacks basic evidence that I possess. So it wouldn't perhaps be too surprising if he, unlike me, failed to have a justified belief about the weather.

So I no longer consider respecting the New Evil Demon intuition in generality a central desideratum for a theory of justification. I continue to recognise it as somewhat intuitive, but I now categorise it with other internalist intuitions that I am comfortable rejecting if necessary. I certainly don't think it's obviously *false*—and although the theory of justification I'll go on to develop here won't *entail* it, it is consistent with it. But I still consider many internalist intuitions weaker than the New Evil Demon intuition, such as the intuition that Adam's belief's justification doesn't depend on whether the testimony *really was* true, important.

Consideration of reliabilism also points to another criterion for a theory of justification. Some critics have argued that reliabilism is inconsistent with a certain kind of *normative* role for justification.[16] We want a theory of justification to explain why—or at least to be consistent with the apparent fact that—justification is a credit to a believer.

To summarise the discussion thus far, I have articulated four desiderata for a theory of epistemic justification:

1. It should not be too stingy; at least many of our ordinary beliefs should be counted as justified.
2. It should be comprehensible and illuminating; it should be offered in terms of notions we understand, which improve our understanding of justification.
3. It should respect at least some internalist intuitions; it shouldn't make it too easy for changes in the external environment, which make for no differences in a subject's experience, to make for a change in the justification of the subject's beliefs.

[15] E.g. Goldman (1988), Engel (1992), Miracchi (2016), and Sosa's (pp. 159–61) contribution to (BonJour and Sosa, 2003).

[16] See e.g. Zagzebski (1996, pp. 300–3).

4. It should explain or illuminate why justification is a credit to the believer.

A theory that meets all of these criteria would be worthy of serious consideration—few if any of the views in the extant literature do so. Note that I do *not* include on my list of desiderata that a theory of justification should be given in 'non-epistemic' terms. I see no particular motivation to respect Alvin Goldman's constraint, articulated thus:

> I want a set of substantive conditions that specify when a belief is justified. Compare the moral term 'right'. ... Normative ethics tries to specify non-ethical conditions that determine when an action is right. A familiar example is act-utilitarianism.... [Its] necessary and sufficient conditions clearly involve no ethical notions. Analogously, I want a theory of justified belief to specify in non-epistemic terms when a belief is justified. (Goldman, 1979, p. 1)

I do not attempt to reduce the epistemic to the non-epistemic; I do not consider it obvious that it is possible to do so.[17] It is consistent with my 'relative priority' interpretation of the knowledge first program that some sort of reduction might ultimately be possible—perhaps, for instance, knowledge is ultimately explicable in physicalist terms—but this is not among my commitments. Nor is it circular to characterise justification in terms of knowledge, once we reject the traditional program of attempting to characterise the latter in terms of the former.

4.3 J=K?

With this background in place, I am now prepared to develop my positive proposal. The starting point is this: one dimension of evaluation for beliefs concerns whether they achieve knowledge. A knowledgable belief is in an important sense epistemically good, and an ignorant one is in a similar sense epistemically bad.[18]

One possible view embracing the knowledge-first approach would stop here: a belief is epistemically good just in case it is knowledge. Since justification just *is* the epistemic goodness of a belief, a belief is justified if and only if it is knowledge:

> **J=K** S's belief that p is justified if and only if S knows that p.

This view has been embraced by Sutton (2007) and Littlejohn (2014).[19] It is an instance of a knowledge-first approach to justification, and it performs reasonably well with respect to some of the criteria indicated in the previous section—for example, it is illuminating in the sense that it characterises the explanandum in terms of a notion we understand well, viz. knowledge.[20] If anything like the knowledge first approach is correct, knowledge is just the sort of property we should expect to be highly eligible to play theoretical roles in one's approach to justification.

How does J=K fare with respect to the desiderata of the previous section? Given a plausibly antiskeptical approach to knowledge, J=K also avoids the radically stingy consequences of classical foundationalism. Nevertheless, it's hard to deny that J=K has some surprisingly stingy consequences; many of our intuitive verdicts about justification predict a more liberal theory of justification that J=K. Notice for instance that since there are no false knowledgable beliefs, J=K has it that

[17] Indeed, one of the central conclusions of Ichikawa and Jarvis (2013) is that (for reasons independent of the issues discussed here) it is impossible to do so; Ch. 5 of that book argues that some epistemic properties are primitive.

[18] This leaves open whether there are other sources of epistemic value as well; one needn't say that all knowledgable beliefs are on an epistemic par.

[19] See also Williamson (2016). Kelp (2016) describes views on which J entails K as 'Strong Knowledge First Epistemology' approaches to justification.

[20] See e.g. Nagel (2013) for a summary of empirical evidence about prephilosophical understanding of knowledge.

there are no false justified beliefs. But it is plausible that there are some justified false beliefs—when we respond properly to sufficiently misleading evidence, our beliefs are justified, but mistaken. For example, suppose that Robin exhibits all of the outward signs of being a simple peasant farmer, and that on this basis, Rose believes that Robin is a simple peasant farmer. We can suppose that her evidence for this claim is very strong and she has no particular reason to doubt it. If it turns out that as a matter of fact, Robin is an evil baronet in disguise, it is then extremely natural to call Rose's an instance of a (false) justified belief.

Again, I recognise that the relevant notion of 'justified belief' is an epistemologist's technical notion. In calling the idea that a particular belief is justified *intuitive*, I am *not* making a prediction about what ordinary speaker's natural inclinations would be if you asked them whether the Rose's belief is epistemically justified. (I think the best prediction would be that most of them would just stare.) But there *is* a robust community practice of philosophical discourse in this domain; it's not like we're in a position of deferring to a particular individual about the sense of the term.[21] Note also that it is a consequence of J=K that the epistemological consensus about Gettier cases—namely, that they constitute justified true belief without knowledge—is incorrect. If J=K is correct, Edmund Gettier was making a rather basic mistake in attempting to give examples of justified true belief that do not amount to knowledge; the trivial answer to his eponymous (1963) question, 'Is Justified True Belief Knowledge?', is 'yes (but that's a redundant way of putting it)'.[22]

It barely needs mentioning that J=K performs abysmally with respect to broader internalist intuitions about justification.

Defenders of J=K like Sutton, Littlejohn, and Williamson tend to emphasise the fundamental normative roles they take knowledge to play. For example, Sutton (2007, pp. 44–48) argues that considerations having to do with assertion require that there be a knowledge norm of belief to the effect that one must believe something only if one knows it. I agree that knowledge has important normative roles to play—indeed, I will argue in Ch. 7 that something very much like this norm does apply. So I also agree with Sutton that there is *a sense* in which any belief that fails to be knowledge is thereby defective; there is *a sense* in which any belief that fails to be knowledge is a belief that a subject isn't *permitted* to hold. But even granting all of this, I think the moral to be drawn from this is that the relevant notion of permission is a poor candidate for the normative status to which 'epistemic justification' refers. One can agree with something very much like the normative claims of Sutton and Littlejohn without following them in expressing them in terms of justification.[23]

Is this a mere terminological dispute? To a considerable degree, I suspect it is—their externalist normative position might be much more easily swallowed by mainstream epistemologists if it were couched in less revolutionary vocabulary. But there may well be a substantive dispute as well. For it's not clear that the J=K theorists tend to recognise *any* status that plays the role I say justification does. They often write as if they wish to explain more traditional intuitions about justification as intuitions about blamelessness or excusability.[24] One should certainly recognise the distinction between satisfying a norm and deserving no blame for failure to satisfy a norm. But I do not think

[21] The situation is probably like this with *some* philosophical terms. Perhaps Crispin Wright has stipulative authority over the term 'warrant transmission'. Perhaps Miranda Fricker has stipulative authority over 'hermeneutical injustice'. (See Wright (2000) and Fricker (2009, Ch. 7) respectively.) But 'epistemic justification' gets its sense from the broader community of use. Notice that I am here in implicit disagreement with Cappelen (2012) who argues in a different case (that of 'intuition') that there isn't enough to fix a community meaning because there's no consensus on theory or agreement about enough of the paradigms. The former condition, I think, is unnecessary for community meeting; the latter condition is plausibly satisfied (for both 'intuition' and 'justification'). And indeed, examples of justified false beliefs are often among the paradigms we use in ostending the notion of justification, for example to our epistemology students.

[22] Kelp (2016) gives this complaint against J=K.

[23] Sutton takes himself to have argued that there is *no* normative role for a notion of justification other than knowledge. Sutton (2007, pp. 2–3, 40). I agree with Kelp (2011) that there's not really much of a case for such a strong claim.

[24] See e.g. Williamson (2016).

that their recognition of the latter category suffices for recognition of (what I think epistemologists really do call) justification. There is a difference, for example, between the aforementioned case of Rose, who is duped into a justified belief by a skilled deception, and the corresponding case of Zorah, who forms a false belief because, due to her limited experience and extreme naivety, she accepts the testimony of obviously untrustworthy people, even in the face of overwhelming evidence to the contrary. Suppose, for example, that Dick tells Zorah that he'll be faithful to her forever, and that she believes him right away, despite his long track record of feckless transfers of affections. If Zorah doesn't have enough experience to recognise the signs of unreliable testimony, we may well hold her *blameless* for her mistaken belief, *even while* judging it to be unjustified. Her failure is importantly different from Rose's, the result of a genuine epistemic skill that happens to have misfired in an uncooperative environment. So I think it would be a mistake to assimilate justification to either knowledge or to blameless belief.

Ultimately, the problem with J=K is that it *collapses a distinction*. Epistemic discourse recognises distinct categories of knowledge and justified belief; to argue that this is an error would require arguing, not only that *some* norms do not distinguish the two, but that *none* do so.[25] Such a universal claim carries a significant burden; I do not see that Sutton, Littlejohn, or Williamson have met it. Rather than refusing to recognise an additional category that corresponds to traditional philosophical judgments about intuition, then, it would be more prudent to allow that there is a distinct category of justification—even if we are as-yet agnostic with respect to how theoretically interesting it will ultimately prove to be. Many epistemologists have thought that justification will play central roles in the normative parts of epistemology; to accept the knowledge first program is to think that knowledge will play many of the most central roles—this does not require legislating in advance that no distinct notion of justification may also play significant roles.

But J=K does get something right: we *can* use knowledge to make sense of justification, reversing the traditional order of explanation. I'll try to offer a different view—one that does less intuitive violence to intuitions about justified beliefs—that retains this feature.

4.4 Justification as Potential Knowledge

Here is the key idea motivating my approach to justification. Consider again cases of false justified belief like Rose's. The key positive status that Rose's belief seems to have is that the reason it fails to be knowledge has to do with features of her external environment outside of her awareness. She processed all her evidence in just the kind of way that one might have expected would lead to knowledge—the reason it didn't lead to knowledge is that the broader external world didn't cooperate. In particular, her interlocutor was lying.

We might codify this idea into a theoretical principle like this one: What makes a given belief justified is that, if it fails to be knowledge, that is due to features of the world outside of the subject. Or equivalently, what makes a given belief justified is that the subject's own state is consistent with that belief's being knowledge. Let's consider

> **JPK** S's belief is justified iff there is a possible individual, alike to S with respect to all relevant basic evidence and cognitive processing, whose corresponding belief is knowledge.

Obviously 'relevant' here is a term that will want some further unpacking; we'll get to that later. I call the view 'JPK' for 'Justification is Potential Knowledge'. More colloquially, the idea here is

[25] Compare similar remarks in Kelp (2011, pp. 393–5) and Ichikawa (2014, p. 188). Jessica Brown also gave arguments of this kind in her (2016) APA symposium presentation.

that the subject *herself* has done enough so that, if the rest of the world cooperates, the belief will be knowledge.

In Ichikawa (2014) I defended a more specific internalist thesis:

> **JPK**$_i$ S's belief is justified iff there is a possible individual, alike to S with respect to relevant intrinsic states and cognitive processing, whose corresponding belief is knowledge.[26]

Notice that JPK$_i$ trivially satisfies the New Evil Demon intuition; it directly builds in the idea that external differences can't themselves make for a difference in justification. For the reasons discussed in §§3.11–3.12, I no longer consider respecting such a strong internalist intuition to be a central motivation worth aiming for. My present generalization of JPK$_i$ into JPK is a way of being neutral on this matter. If the 'relevant respects' are all exclusively intrinsic, then JPK$_i$ is equivalent to JPK. But given the disjunctivist neo-Moorean position gestured at in the previous chapter, we may also wish to be open to factive perceptual states as among the relevant respects. If so, then JPK will not entail JPK$_i$. It is possible, for instance, for two intrinsically identical subjects to differ with respect to their factive perceptual states; JPK allows that this may suffice for a difference in justification. If Despard sees that a villager is sick, and Margaret has an intrinsically identical hallucination, Despard's belief may be justified (indeed, it is knowledge), even though Margaret's is not (since her basic-evidence counterparts are hallucinating).[27] So I am now more attracted to this *perception*-emphasising version of JPK:

> **JPK**$_p$ S's belief is justified iff there is a possible individual, alike to S with respect to all relevant factive perceptual states and cognitive processing, whose corresponding belief is knowledge.

For much of this chapter however, I'll focus on the more general claim of JPK; I don't think the neo-Moorean disjunctivism plays any role in making JPK more plausible than competing theories of justification.

4.5 Is JPK Internalist?

Notoriously, epistemologists use 'internalism' as a name for at least two importantly different ideas. A *supervenience* sense to internalism has it that justification supervenes on certain features of the

[26] In that paper I called this view 'JPK'; I have elected here to use that name for the more general view, adding the subscript now for clarity. I also stated the view then in terms of 'counterparts' rather than 'possible individuals'; I now prefer the latter for coming with less theoretical baggage. (JPK does not depend on substantive questions about which possible individuals are different versions *of me*.) I also spoke then simply of being 'alike in all relevant intrinsic respects', rather than separating out the basic evidence and cognitive processing—I had previously been supposing that each was among the intrinsic; but the distinction is more important for the general version of the view, so I include it here. If the view of my previous paper is right, the distinction is idle.

[27] In fact, the disjunctivist version of JPK doesn't *require* that hallucinatory beliefs are unjustified, although it fits rather naturally with that idea. The reason it *could* allow justification is that although Margaret's basic evidence does not include the proposition that a villager is sick, it is arguably some evidence in favour of that proposition. Suppose that Margaret's relevant basic evidence includes only that she has certain kinds of experiences—experiences as of a sick villager. Suppose also that it is possible for someone to know that a villager is sick on the basis of such experiences. (A core commitment of the disjunctivist approach is that this is not the *usual* route to perceptual knowledge, but it's not clear that it must be ruled out as impossible.) Then consider a possible individual alike to Margaret with respect to basic (intrinsic) knowledge, who then comes to know (indirectly, via inference on her perceptions) that there is a sick villager. So it looks like it may be possible for a disjunctivist JPK theorist to embrace the supervenience of justification on the internal. Some of my other commitments, however, may rule this out. (Note that given my neo-Lewisian contextualism, the possibility that Margaret's senses do not match reality is relevant by the rule of actuality; no one can be said to 'know' on the basis of sensory experiences that a villager is sick. I'll explain the role of contextualism in my approach to justification in §4.6.)

subject's state. This can come in various forms—a *geographical* supervenience thesis might have it that justification facts supervene on the facts inside a subject's head. A *mentalist* supervenience thesis would have it that justification facts supervene on the subject's mental states.[28] The latter doesn't entail the former, since mental states are individuated broadly; so some internalists strengthen the mentalist claim, saying that justification supervenes on *nonfactive* mental states.[29] JPK is straightforwardly a supervenience thesis of this broad type: it has it that there can be no difference in justification facts without a corresponding difference in the subject's relevant situation. Although I do not (any more) assume that all such relevant features are geographically internal—I allow that factive perceptual states are relevant features—JPK is still *mentalist internalist* approach to justification; differences in the external world that do *not* make a difference with respect to a subject's mental states do not affect whether a belief is justified. The relevant facts comprise the subject's basic evidence and cognitive processing.

Another sense of internalism departs from a supervenience claim and has it instead that a subject has a kind of privileged, first-personal access to facts about justification or the features that justify.[30] I've never been very clear about how if at all one can make sense of this kind of *access* internalism.[31] I make no claims as to whether JPK is internalist in this access sense, since I don't understand it.

I expect that some internalists will complain that JPK_p is not internalist in *any* sense. (I expect that everyone will admit that JPK_i is.) One might insist that *any* form of internalism must respect the New Evil Demon intuition in generality; but JPK_p doesn't do that. I'm not very interested in arguing about just how we should or shouldn't use the word 'internalism'; I don't particularly care whether other epistemologists follow me in using it in a broad enough way to encompass JPK_p. But I would like at least to explain why I think it's a reasonable application. To my mind, the core idea of the supervenience sense of internalism is that justification is a matter of getting things right, from an epistemic point of view, within a particular local sphere. To repurpose a metaphor Timothy Williamson has used in a different way, justification is a matter of keeping one's 'cognitive home' in order. (I'll discuss this metaphor further in §4.9 below.) Just what the domain of that sphere is, different internalists may disagree about—some think, for example, that only conscious-to-mind experiences count, while others hold that stored memories are in the relevant sense internal. I'm inclined to think that the contents of factive perceptual states are also in the relevant sense internal. What all of these views have in common is that they deny that factors outside of a subject's mental sphere have any bearing on justification. So I will continue to describe JPK as an internalist view.

Unusually for internalist views, JPK retains a deep connection between justification and knowledge. According to JPK, a belief is justified if and only if it is sufficiently similar to an instance of knowledge; the respects of similarity at issue are the relevant mental ones.

The hedge 'relevant' is to rule out demanding that the possible knowers be alike to the justified believer with respect to features concerning the content of the belief, but irrelevant to its cognitive realisation—for example, JPK should not require that my relevant counterparts' brains are physical duplicates of mine; if it did, JPK would have the implication that I cannot have false beliefs about my brain.[32] The idea behind JPK is that a subject's belief is justified just in case her basic evidence

[28] Feldman and Conee (2001).

[29] Wedgwood (2002) is an example.

[30] For example, Carl Ginet writes that '[e]very one of every set of facts about S's position that minimally suffices to make S, at a given time, justified in being confident that p must be directly recognizable to S at that time.' Ginet (1975, p. 34) See (Pappas, 2014, §§2–3) for discussion and further references.

[31] Reliabilism is supposed to be paradigmatic of a view that fails to be access internalist, but reliabilists can certainly allow that subjects know what evidence they have, or whether their beliefs are justified. See also §3.11, where I canvass a few reasons to be skeptical of the idea that we have a kind of privileged access to features of one's own subjective experience.

[32] Suppose, for example, that I have misleading evidence suggesting that I have a concussion. On the interpretation of JPK dismissed here, assuming that having a concussion impacts one's mental situation, I could not have a justified belief that I have a concussion, as none of perfect mental duplicates has a concussion, and so none of them can know that he has a

and total cognitive processing are consistent with her having knowledge. It is trivial on JPK that all cases of knowledge are cases of justified belief. So JPK captures the right-to-left direction of J=K, and thus performs at least tolerably with respect to the 'not-too-stingy' desideratum, on the hypothesis that many people have a decent amount of knowledge. But it does not accept the left-to-right direction, which ran so starkly against the internalist intuitions. Subjects in unfavourable epistemic circumstances—like Rose who has misleading evidence about Robin's status—are among those who have justified beliefs that fall short of knowledge. There is a possible person just like Rose internally—with the same basic evidence (both her factive perceptual states and subjective experiences), and who processes her evidence in just the same way, who ends up with a knowledgable belief.[33] Since Rose's relevant internal situation is consistent with her knowing, her belief is justified. So JPK respects the moderate forms of the internalist intuitions mentioned in §4.2.[34]

Similarly, JPK delivers the intuitive verdict that subjects in Gettier cases have justified beliefs—had the external world cooperated more fully, their internal states would have realised knowledge. JPK also meets the illumination desideratum, as it does not posit any mysterious *sui generis* states or relations. JPK does not imply that there can be no justified beliefs in necessary falsehoods, because the relevant corresponding belief needn't be a belief with the same content. I'll explain and explore this flexibility in the Appendix to this chapter.

The extension JPK offers to justification is reasonably intuitive. As just indicated, many ordinary perceptual beliefs, including some that fall short of knowledge, are justified. But JPK also places plausible constraints on justification: intuitively unjustified beliefs are typically such that the subject's internal state precludes knowledge. For example, if I flip a fair coin and confidently predict on the basis of a mere special feeling that it will land tails, my belief is intuitively unjustified, whether or not it is true. The present view delivers that verdict, as anyone alike to me with respect to my basic evidence and cognitive processing would be forming beliefs irresponsibly in a way inconsistent with their coming to knowledge, however the coin may land.

4.6 Contextualism

In the previous section I stated JPK in the object language. Since I am a contextualist about 'knowledge', I think the right-hand side of the statement is context-sensitive:

> **JPK** S's belief is justified iff there is a possible individual, alike to S with respect to all relevant basic evidence and cognitive processing, whose corresponding belief is knowledge.

As in the case of evidence, there are various options for how to interpret this link. And again as in the case of evidence, I prefer a version of the view that holds the statement above in metalinguistic generality, along with a corresponding contextualism about 'S's belief is justified'.[35] That is:

---

concussion. Similarity with respect to whether one has a concussion is not relevant in the sense of JPK.

[33] I here assume that Rose herself doesn't have, among her basic evidence, *that the person before her is a peasant farmer*—although she can directly perceive some external-world propositions (perhaps she can directly perceive *how he is dressed*, and *that he says he's a peasant farmer*), she can't directly perceive such specific ones. There are stronger and weaker possible neo-Moorean views in this neighbourhood.

[34] *Radical* skeptical scenarios, where Rose fails to have the factive perceptual states she actually has, are a different story. Delivering the same story about these cases is what would be required for the New Evil Demon intuition in generality; JPK_i vindicates this stronger claim, but JPK_p may not.

[35] Contextualism about 'justified belief' has been defended by Stewart Cohen and Ralph Wedgwood. Their views differ from mine in important ways. Wedgwood's version of contextualism is explicitly not one that is based on a Lewisian relevant alternatives theory (Wedgwood, 2008, p. 14); he also doesn't seem interested in connecting his view about justification ascriptions to contextualism about knowledge ascriptions. Cohen's approach is more similar to mine in some respects—he

> **JPK**$_M$ For any context C, 'S's belief is justified' is true in C iff there is a possible individual, alike to S with respect to all relevant basic evidence and cognitive processing, whose corresponding belief satisfies 'knows' in C.[36]

JPK$_M$ is the natural interpretation of JPK within the contextualist framework of this book; just as there are knowledge ascriptions that are true in non-skeptical contexts and false in skeptical ones, so too will 'justified belief' ascriptions vary in truth conditions according to conversational contexts. In contexts in which the possibility that one's conversational partner is being deceptive is relevant, for instance, where we will say that one 'doesn't know', we'll also say that one 'isn't justified in believing'. Testimonial cases don't become Gettier cases when the standards are raised.

In fact, this contextualist approach to justification can helpfully accommodate certain counterorthodox intuitions about Gettier cases. In my introductory epistemology courses, when I teach Gettier cases for the first time, I inevitably get one or two students who wish to deny that the subjects have a justified belief. If, for instance, I discuss a 'stopped clock' case[37], where orthodoxy has it that one is justified in one's belief about the time on the basis of the clock's reading, some students will argue that this is mistaken—one should have checked to make sure the clock was working, before forming beliefs about the time on its basis. It is usually not difficult to talk them down from that verdict—'do you think you should *never* trust a clock before making sure it is working? How often do you really check this kind of thing?'—but the initial reluctance to ascribe justification is genuine. On a contextualist approach like the one I'm defending here, this is easily explained. Mention of the fact that the clock is stopped tends to create a skeptical conversational context in which that possibility is relevant. So my students who insist that the Gettier beliefs aren't justified are focused a higher epistemic standard than the one I intended for them. By shifting their attention to more ordinary cases, I contribute to lowering the standards again, in order to demonstrate how knowledge can come apart from justified true belief.

For the most part, contextualism about justified belief won't play a central role in the discussion or defence of JPK. (Indeed, my first presentation of the view, in Ichikawa (2014), didn't discuss contextualism at all.) In general, the context will make it clear how demanding an epistemic standard we're working with at a given time, and I won't focus on particular puzzles that contextualism will resolve.[38] So most of the chapter will be written simply in the object language; I'll ascend semantically only on occasion.

---

observes that justification ascriptions are consistent with there being some possibility in which the grounds hold and the justified belief is false, and that context determines 'how good one's reasons have to be'. Although he doesn't explicitly put this notion in terms of ruling out alternatives, that would be a reasonably natural development of his approach. Unlike Cohen, however, I do not think that the context-sensitivity of 'justified belief' can support an *argument* for contextualism about 'knows'; Cohen suggests that '[s]ince justification is a component of knowledge, an ascription of knowledge involves an ascription of justification' and so will be context-sensitive. Cohen (1999, p. 60). But I deny that ascriptions of knowledge 'involve' in any interesting sense ascriptions of justification. (And even if they did, inferring contextualism about 'knows' from contextualism about 'justified' would arguably amount to a use-mention confusion of the kind described in my Introduction.)

[36] In fact, given the contextualist approach to evidence ascriptions defended in Ch. 3, one could also invoke the shiftiness of 'evidence', dropping the reference to *basic* evidence, thus: 'S's belief is justified' in C iff 'there is a possible individual alike to S with respect to all evidence and cognitive processing, who has knowledge' is true in C. I consider the version in the main text to be easier for the purpose of theorizing, as it has fewer 'moving parts'.

[37] Russell (2009 [1948], pp. 170–1)

[38] An anonymous referee wonders whether, in a way similar to what I suggested of about 'knows' and 'evidence' in §3.8, 'knows' and 'has a justified belief' may have divergent pragmatic effects on the conversational score. Perhaps, for instance, talk of justification tends to shift speakers into more skeptical contexts. This strikes me as an interesting idea, worthy of further exploration.

4.7 Steven Reynolds

Steven Reynolds has developed a view similar to JPK. Reynolds's view, first sketched in Reynolds (1998, pp. 533–6) and developed in Reynolds (2013), is that S's belief is justified just in case it has, for S, the *appearance of knowledge*. This view has many similarities to mine—it explains justification in terms of knowledge; more specifically, it encodes the idea that a justified belief that fails to be knowledge is in some way deviant or unexpected. It is also, like mine, a serious attempt to capture what's worth capturing about internalist intuitions about justification.

Prima facie, there appear to be clear counterexamples to the thesis that the appearance of knowledge is necessary and sufficient for knowledge. Against the necessity of this appearance for knowledge, one might suppose that subjects with no concept of knowledge, and who therefore are not susceptible to the appearance of knowledge, might nevertheless have some justified beliefs. Against the sufficiency of this appearance for knowledge, one might consider cases of intuitively unjustified dogmatic belief, where the appearance of knowledge is present (though it shouldn't be). In response to these counterexamples, Reynolds clarifies that he has a more nuanced understanding of the notion of 'appearance'. He emphasises, for instance, that one may have the appearance of something without any conception of that thing—he gives the example of a birder who observes an unfamiliar bird that happens to be a Stellar's Jay.[39] Reynolds's idea is that the birder *does* have *the appearance of a Stellar's Jay*, even though it would not be correct to say that it *appears to the birder* as if there is a Stellar's Jay. In the same way, subjects who lack the concept KNOWS may nevertheless enjoy the appearance of knowledge. As for subjects who wrongly take themselves to know, Reynolds denies that, in the sense he intends, they enjoy the *appearance of knowledge*. (Perhaps it *appears to them as if* they know, but that's a different matter.) Subjects can be mistaken about what appearances they're undergoing, and that's what Reynolds thinks is happening when one has an unjustified dogmatic belief—they *wrongly take themselves* to have the appearance of knowledge.

I confess that I find the notion of appearance appealed to somewhat obscure. When I see Fred (whom I know nothing about), do I have, in the relevant sense, the appearance of his twin brother George? Does my twin who is deceived by an evil demon have the same appearances I do? Reynold's discussion of appearance isn't thick enough to help me see just what the relevant notion amounts to. One way of developing the view would have it that one has the appearance of X just in case one has the basic evidence one would have if one knew that X. This would bring Reynolds's view closely in line with JPK. Perhaps there are more ways of developing substantive views of the relevant notion of appearance, but at present, I don't see an obvious case for considering the best version of the approach to be in competition with JPK.

4.8 Alexander Bird

Alexander Bird has also offered an account of justification that is similar to JPK. Here is Bird's view:

> **JuJu** If in world w_1 S has mental states M and then forms a judgment, that judgment is justified if and only if there is some world w_2 where, with the same mental states M, S forms a corresponding judgment and that judgment yields knowledge. Bird (2007, p. 84)

Bird's view is like mine in that it is one according to which, in his words, 'justified judging is a certain kind of approximation to knowledge.' (p. 85) (Note that it would be straightforward to implement JuJu in a contextualist framework similar to mine if one wished to do so; I'll leave this tacit

[39] Reynolds (2013, p. 369)

here.) The main difference between Bird's view and mine is that his requires sameness of mental states generally; JPK's requirement is for sameness of basic evidence and cognitive processing.

Because he requires *all* mental states to be held in common, Bird needs a diachronic version of the view. Since he agrees with Williamson (2000) that knowledge is a mental state, if Bird required a subjects with justified judgments to have *all* mental states in common with subjects with knowledge, even after forming their respective beliefs, JuJu would entail the problematic left-to-right direction of J=K. After all, since knowledge is a mental state, anyone with the same mental states as someone who knows that p must herself know that p. Bird avoids this implication by declining to require that *after the judgment is made*, one has the same mental states as a possible knower.

But this diachronic feature of Bird's view introduces challenges not shared by JPK. Since the relevant constraint only concerns mental states *prior* to the judgment, for example, Bird's view cannot allow that *concurrent* or *future* mental states can bear on justification. But it is not at all clear that a theory of justification should rule this out; suppose Roderic has total mental state M at time t_1 in worlds w_1 and w_2. In each world, Roderic goes on to form the judgment that p at subsequent t_2. Suppose that in w_1 this judgment achieves knowledge, but in t_2, it is false. Bird's view has it that in w_2, Roderic's belief is justified, *regardless of what other judgments* he makes at t_2. But it is consistent with these stipulations that in w_2, (but not in w_1), Roderic went on to form *other* beliefs at t_2 that were in tension with p. Suppose for instance that in w_1 Roderic hears a declaration of love from Hannah and comes to know on this basis that Hannah is in love with him. In w_2, things start the same way, but at the same time as he forms the belief that Hannah is in love with him, he *also* forms the paranoid belief that Hannah is attempting to deceive him for some nefarious purpose. This latter belief, we may suppose, is wholly unjustified—it is a thoroughly irrational, albeit genuine belief. And bizarrely, this belief doesn't stop him from believing that Hannah is in love with him on the basis of her testimony; in w_2 he forms both beliefs together. While I don't think it's *obvious* what to say about such an admittedly odd case, it seems to me at least reasonably plausible that Roderic's paranoid belief serves as a defeater for the belief that Hannah loves him; that therefore, in w_2, Roderic's belief that Hannah loves him is unjustified. But Bird's diachronic approach cannot allow this, since everything prior to t_2 is the same in both worlds, one of which involves knowledge. By contrast, JPK does permit this verdict, by requiring the sameness of all relevant states even at t_2, the time in which he has the belief in question. If Roderic has new beliefs that undermine his justification, then his twin too has new beliefs, which plausibly undermine his knowledge.[40]

Another potential concern about Bird's view derives from content externalism. The contents of beliefs depend in part on one's external environment. As a result, requiring sameness of mental states places nontrivial constraints on the external environment of the possible knower. For example, anyone with beliefs about Hesperus is limited in mental-state duplicates to subjects thinking about the object that is in fact Venus. This feature of Bird's view, like its diachronic nature, derives from emphasising sameness of *mental state* in particular. And as in the previous case, it leads to counterintuitive consequences.

Suppose that someone believes on the basis of misleading evidence that Hesperus is not Phosphorus. Bird's view does not straightforwardly entail that this belief cannot be justified, since it may be based on antecedent mental states that are consistent with having come to a knowledgable belief, had the external world been sufficiently different. But it does entail that no *subsequent* beliefs based on this belief can be justified. For suppose one is in this position, and also learns that Hesperus is Venus. If she inferred that Phosphorus is not Venus on the grounds that no two distinct bodies can both be Venus, then on plausible assumptions, Bird's view has it that this belief cannot

[40] As Jessica Brown pointed out to me, a natural move on Bird's behalf would be to reach for a 'no defeaters' clause here to add to his account; this could rule out the kind of case I am considering. But defeaters are good candidates for epistemic properties that must themselves be understood in terms of justification; the challenge for one wishing to invoke defeaters at this stage would be to give an independent characterization of them.

be justified. For Bird's view is that it is justified only if a possible subject with the same antecedent mental state comes to knowledge. But the antecedent mental state includes the belief that Hesperus is not Phosphorus, which is necessarily false. If, as orthodoxy has it, inference from a false belief is inconsistent with coming to knowledge, no one with those antecedent mental states will come to know.[41]

JPK, too, faces important questions about how it handles justified beliefs in necessary falsehoods—I'll consider them, and highlight the comparison with Bird, in the Appendix to this chapter.

A third possible problem for Bird's view is that it doesn't seem general enough. Because of its diachronic nature, the approach is naturally given in terms of *judgments*, rather than beliefs. It's easy to think this is not a very substantive issue, as there is a systematic relationship between judgment and belief: a judgment is the formation of a belief. And perhaps, one might think, a belief is justified if and only if it is the product of a justified judgment.[42] But I do not think it's obvious that things are this simple. Suppose for instance that some people have some *innate knowledge*. If so, then they have justified beliefs (since all knowledge is justified) that were never formed by a judgment (since they are innate). But Bird's approach cannot explain the justification for any such beliefs; it only applies to judgments. Then again, it also seems as if the product of a justified judgment might later be defeated by the acquisition of new evidence; if the belief is nevertheless inappropriately retained, it should be counted as unjustified; but Bird's view has no resources to say this—it only applies to the *formation* of beliefs.[43]

So although Bird's view is very similar to JPK in spirit, I believe that JPK is preferable in detail.

4.9 Justification as a Normative Status

In setting out the desiderata for a theory of justification in §4.2, I laid out four criteria: a theory of justification shouldn't be too stingy, it should be illuminating, it should be respectful of at least many internalist intuitions, and it should explain why justification is a credit to the believer. So far, I hope to have made a case for the first three of these desiderata. What of the fourth?

One objection I have sometimes heard to JPK targets just this line.[44] The objection runs thus: According to JPK, being justified isn't a matter of what the subject herself is doing, epistemically speaking—instead, it considers whether some *other* possible person, in a *different* possible world, has knowledge. But (the objection goes) if we're assessing *this* individual, why should we care about whether some *other* person has knowledge? I have never been much moved by this worry. I think it's just a misdescription of the view to suggest that JPK asks one, when considering whether Richard's belief is justified, to focus instead on some other possible individual and assess him for knowledge. JPK has it that Richard's belief is justified just in case *his own* basic evidence and cognitive processing are consistent with the resultant belief's amounting to knowledge. Whether he has a relevantly similar counterpart of the right sort is wholly a matter of how he himself is.[45]

[41] This orthodoxy is not without dissent. See e.g. Warfield (2005), Fitelson (2010), Luzzi (2010), and Luzzi (2011). I will discuss and resist some of the reasons for dissent in §5.8.

[42] This seems to be Bird's own attitude; he often writes of his view as if it concerns the justification of 'judgements and beliefs'; on p. 85 he remarks that we 'think a belief is justified because of the way in which it came to be formed'.

[43] A fourth potential challenge for Bird is that his approach fails to respect the internalist motivations for which he aims. See McGlynn (2012, pp. 364–5) for an argument to that effect. I do not emphasise this challenge in the main text here (though I did in Ichikawa (2014)) because it is, as McGlynn points out, a sensitive matter to determine just what the relevant internalist intuitions *are*, and some ways of developing JPK may not fare much better than Bird's view with respecting at least some of the relevant internalist intuitions. Although JPK_i, for instance, delivers the 'new evil demon' intuition, JPK_p might not. (But see fn. 27.)

[44] I first heard this objection from Allan Hazlett, who raised at a talk I gave on this material at the University of Edinburgh. It was also pressed to me by a referee for the *Canadian Journal of Philosophy*, where I first published a version of these ideas.

[45] The worry is reminiscent of Saul Kripke's 'Humphrey' complaint against David Lewis's counterpart theory; Humphrey

The central idea to JPK is that justification is a kind of approximation of knowledge; it says that a belief is justified if and only if a certain core part of the subject's cognitive system is such that it could produce knowledge, if the rest of the world would only cooperate in the right way. JPK is a version of the idea that Martin Smith sets out in broad outline in this passage:

> Justification, I take it, is what makes a belief a good candidate for knowledge. I'm happy to leave this suggestion somewhat imprecise for the time being, but the basic idea is this: My belief is justified just in case I have *done my epistemic bit*—the rest, as it were, is up to fate. My belief will qualify as knowledge provided that the world obliges or cooperates—but *I* am not required to do anything further. Smith (2010, p. 12)

I think this idea gets to the heart of the internalist intuitions about justification. Many internalist ideas have faced very strong challenges in recent decades, some of which I discussed in Ch. 3. In particular, I am thoroughly convinced, in part by the arguments of Williamson (2000, Ch. 4), that there is no special internal domain that is a "cognitive home" in the sense that within it, one can never be wrong about whether a feature obtains, or even only that when a feature obtains, one is always in a position to know that it does. The idea that there can in some sense be no doubts about the internal was a mistake. Nevertheless, these arguments do *not* establish that there's no *other* interesting sense in which there is a kind of "cognitive home". Smith's framework here suggests the following idea: there *is* a kind of "internal" cognitive home—this doesn't mean we can't be wrong about it or are guaranteed access to it, but rather that we have *special responsibility* for it. If one fails to achieve knowledge because the broader world isn't cooperating, that's one thing—there is a distinctive kind of epistemic error that comes of the mismanagement of one's basic evidence and cognitive processing. The justification norm can be thought of as an injunction to keep one's cognitive home in order.

I used scare quotes for "internal" in the previous paragraph because I was working at a level of abstraction that is neutral about just which elements are part of a subject's cognitive home. This is the same choice point I've been mentioning throughout this chapter; as we have seen, JPK is consistent with various ways of developing just what one's cognitive home comprises—in Smith's terminology, in just which things count as among what 'I' must do, and which are a matter of 'fate' or 'the world obliging'. JPK_i can be thought of as identifying the cognitive home with the intrinsic; it encodes the norm that one is in some special way responsible for the intrinsic: one satisfies the justification norm if one is, *on the inside*, just like someone with knowledge. This might be motivated by the idea that justification is a kind of blamelessness, and that the relevant notion of blamelessness is one that refuses to hold subjects responsible for the extrinsic. By contrast, JPK_p includes the contents of factive mental states as part of the cognitive home; it holds that one is responsible for responding properly to that which one sees for oneself. I've been considering these two versions of the view so far, but there are other candidates as well; they correspond to different commitments about just what the normative status of justification governs.[46]

---

cares whether he might've won the election, but he 'could not care less whether someone *else*, no matter how much resembling him, would have been victorious in another possible world.' Kripke (1980, p. 45). But it is in virtue of the way Humphrey himself actually is that other *similar* possible individuals win.

[46] Smith's own view ultimately doesn't provide a determinate answer to the question; his primary focus is on arguing that one's total evidence *nomically supports* a proposition in order for that proposition to be justified; this is unpacked as the requirement that 'the most normal worlds in which E is true are worlds in which P is true'. Smith (2010, p. 17). His broad characterization of justification in terms of knowledge quoted above notwithstanding, it is not at all obvious how to fill out this view into an approach to justification as having done enough 'oneself' for knowledge. One reason this is difficult is that Smith is explicit (p. 26) in signing up only to the necessity, not the sufficiency, of this condition for justification; another is that, although he's not explicit about this, Smith seems to be focused, in discussing his nomic support theory, on

Bird's view can be thought of as the view that all past mental states, including knowledge, make up the cognitive home. In that sense, Bird is correct to claim his view as a 'mentalist' one, but Aidan McGlynn is also plausibly correct to complain that it isn't one that captures many internalist intuitions.[47]

At an extreme, one might hold that *all* knowledge facts, including those that obtain at the time of judgment, are part of one's cognitive home—that one has special first-personal responsibility over all of one's knowledge. On such a view, any belief that falls short of knowledge *for any reason* is a case of failure to keep one's cognitive home in order. That view would be JPK$_k$:

> JPK$_k$ S's belief is justified iff there is a possible individual, alike to S with respect to all knowledge, whose corresponding belief is knowledge.

This is equivalent to J=K.[48] For reasons discussed in §4.3, I think that embracing J=K would be a mistake, but it is useful to observe that it exists as part of a framework for encoding the normative role of justification. If one wished to insist that there is *no* relevant normative epistemic difference between someone who fails to have knowledge because the external world doesn't cooperate, and someone who fails to have knowledge because she processes her evidence irrationally, then J=K might be the most attractive version of JPK. But I think that the more internalist versions I've been considering is preferable for this reason. (JPK$_p$ counts as 'more internalist' than J=K, because it respects more of the internalist intuitions.)

Then again, one needn't necessarily choose one of these notions as the one true theory of justification. As highlighted in the introduction to this chapter, 'justification' is a philosopher's term of art, and I am engaged in a project of explication. There are various notions one can make sense of here. Maybe we ought to retain JPK$_i$, JPK$_p$, *and* JPK$_k$ (J=K), because they *each* encode genuine epistemic norms. I do not rule this situation out, but I am more inclined to suppose that one of them (or an unarticulated competitor) is likely to be theoretically privileged. As in the case of the parallel discussion of evidence, I am inclined to treat JPK$_p$, which emphasises the contents of factive mental states, as the best theory of justification.

4.10 An Objection

JPK has it that, roughly, a subject's belief is justified if her cognitive home is *consistent* with her having knowledge. There is a sense, then, in which the condition for justification is logically weak; my belief is justified if there is *any possible version* of me, who retains the same cognitive home, who knows. This weakness was by design—one of the criteria I began with for a theory of justification was that it shouldn't be too stingy. Still, one might wonder if I have overcorrected for too-stingy views; is the relevant sense of possible knowledge too weak a condition to suffice for justification?

One might be particularly worried about this question in light of content externalism; it is *not* a requirement of JPK that the possible knower know *the same content* as that believed by the justified believer. (See the Appendix to this chapter for more on this feature of JPK.) It's a consequence of my framework that the JPK theorist enjoys considerable flexibility in finding the relevant possible

propositional, rather than on doxastic, justification. (Nomic support is neutral on whether the conclusion is believed. See also McGlynn (2012, p. 365) for the same attribution.) But doxastic justification must be the status he had in mind in the quote above; for if one doesn't believe that p, one hasn't 'done one's epistemic bit' to know that p.

[47] Bird (2007, p. 97); McGlynn (2012, pp. 365–6).

[48] Something like this thought might be interpreted in Williamson's remark that his 'E=K' is a kind of foundationalism, according to which all knowledge is foundational. See Williamson (2000, p. 186). But at the time of *Knowledge and Its Limits*, Williamson did allow for justified false belief.

knowers than might at first be apparent. Does it offer so much flexibility that it might make practically any belief justified?[49] Consider what is intuitively a subject's *deeply irrational* belief. Suppose for instance that Margaret has had a course of experiences relevantly similar to the kinds most people have, but who believes that orangutans are a kind of fruit beverage. The example comes from Tyler Burge, who thinks that in such cases, the beliefs required would be *so* unreasonable that it is generally preferable to suppose that semantic confusion is resulting in a confused utterance, rather than the genuine expression of an irrational belief:

> If a generally competent and reasonable speaker thinks that 'orangutan' applies to a fruit drink, we would be reluctant, and it would unquestionably be misleading, to take his words as revealing that he thinks he has been drinking orangutans for breakfast for the last few weeks. Such total misunderstanding often seems to block literalistic mental content attribution, at least in cases where we are not directly characterizing his mistake. Burge (1979, pp. 90–1)

The worry for JPK I am now considering will exploit the flexibility of mental content to argue that even deeply irrational beliefs like these could end up being justified. So suppose that Margaret says, 'I try to drink orangutans every morning, because they are high in vitamin C'. Can we find a possible individual, similar in the relevant ways to Margaret, who expresses knowledge with her corresponding belief? If we can, then JPK incorrectly predicts that Margaret's belief is justified.

Consider a possible version of Margaret, Margaret′, who inhabits a world where members of her linguistic community use the word 'orangutan' to refer to orange juice. When Margaret′ says to herself, 'I like to drink orangutans', she is expressing the true thought that she likes to drink orange juice. For such a case to be a problem for JPK, we need two more things to be true of Margaret′: her true belief must be knowledge, and she must share Margaret's cognitive home. It's easy to point to *possible* versions of Margaret′ who fail to have both of these features; my task is harder than that. JPK has it that if there is *any possible* person who has both of these features, then Margaret is justified. Still, I think that it's not too hard to show that Margaret′ can *not* have both of these features.

Part of the reason why Margaret's orangutan belief is so irrational is that she seems deeply out of touch with her linguistic community. On the most natural ways of developing the case, Margaret often encounters evidence against the proposition that orangutans are juice. For example, you'd expect that sometimes when she goes to the zoo, she sees an enclosure containing Asian primates marked 'orangutans', even though there is no apparent juice nearby. She has strong evidence against the proposition that orangutans are juice. (If she didn't—if she's in a very impoverished evidential state—then it's not at all clear that her belief *is* deeply irrational. If for instance the only time she's ever encountered the idea of an orangutan was when a prankster told her, 'orangutans are juice', it may be that she has an ordinary justified false testimonial belief.) The problem for the objection is that it is very plausible that this kind of defeating evidence will be reflected in Margaret's cognitive home, and so it, or something suitably like it, will also be present in every candidate for a suitable Margaret′.

For example, JPK_i has it that any suitable Margaret′ must be an intrinsic duplicate of Margaret. But any intrinsic duplicate of Margaret will have experiences that are strong evidence against the idea that orangutans are juice. What those experiences will be will depend on what evidence Margaret has that defeats the justification for her belief, but if, for instance, she's seen signs at the zoo indicating that orangutans are primates, then any suitable Margaret′ will have had sensory experiences as of seeing similar signs, and it is overwhelmingly plausible that this experience would act as a

[49] Thanks to Jessica Brown for pressing a particularly clear version of this worry.

defeater for Margaret′'s justification, just as much as it does for Margaret's. And so, since Margaret′ isn't justified, she doesn't know—which means Margaret isn't justified either. I can't think of any possible evidence that Margaret might have that undermines her justification, that wouldn't come along with intrinsic experiences that would also undermine Margaret′'s knowledge. Very much the same story extends to JPK$_p$. Margaret's cognitive home, on this version of the view, includes visually perceived facts like *that a sign says that orangutans are primates*. So any suitable version of Margaret′ will also have seen that there is such a sign; this evidence will undermine her knowledge.

So I don't see any strong reason for worry that JPK will have implausibly generous implications about justification of deeply irrational beliefs. Such beliefs are irrational for a reason—and that reason, or one suitably similar, will be shared by the relevant counterparts. This shouldn't be surprising. Knowledge does place significant constraints on the believer: it can only occur when one's cognitive home is in order. But a deeply irrational belief just is one in which one's cognitive home is disordered.

4.11 Lotteries

In the final sections of this chapter, I will identify a number of questions on which I am at present officially neutral; various possible answers to them would develop JPK in different directions. This is intended both in order to identify what seem to me to be interesting questions that warrant further study, and to indicate some respects in which JPK can be accepted by epistemologists with diverse commitments.

Contemporary epistemological orthodoxy has it that absent special evidence, owners of lottery tickets that have some small chance of winning cannot know that their tickets will lose.[50] Even if my ticket is in fact a loser, since the evidence against its winning is exactly parallel to the evidence against each ticket's winning, including the winner's, I cannot *know* that my ticket will lose. Within a contextualist framework, it is helpful to ascend to the metalanguage: relative to contexts in which the possibility that A's lottery ticket will win is relevant, 'A knows that her ticket will lose' is false. (Very plausibly, given my Lewisian framework, such sentences will *always* be false when asserted, since this kind of explicit consideration of the possibility will make it relevant.) But relative to less skeptical contexts, such sentences may be true (even if asserting them would destroy such contexts).

Given this orthodoxy, the metalinguistic contextualist version of JPK strongly suggests that unless A has special evidence, 'A has the justified belief that her ticket will lose' will be false relative to the skeptical contexts in which it'd be at home.[51] To put things in the object language, a JPK theorist should hold that one can't have justified beliefs in lottery propositions, because all of the subject's relevant possible counterparts are also in lottery situations, and so can't know.[52] To my mind, this is exactly the right result—particularly given the contextualist apparatus I've already found plenty of reason to reach for, there's good reason to deny both knowledge and justification to lottery beliefs. (In fact, I'll argue in Ch. 7 that we should typically deny *belief* to these cases too.) But JPK itself transcends this approach to both contextualism and the lottery paradox; if one wishes to hold that there can be justified belief and knowledge in lottery cases,[53] JPK does nothing to rule this out. The requirement of JPK is that justification and knowledge be systematically related, but it is silent on the extent of each.[54]

[50] For presentations of this orthodoxy and accompanying discussion, see Kyburg (1961), Nelkin (2000), and Hawthorne (2004).

[51] It doesn't quite *entail* this result, for reasons described in Ichikawa (2014, n. 27), but I think the possible version of the view that rules out knowledge but allows justified belief is implausibly ad hoc.

[52] See Kelp (2014).

[53] Hill and Schechter (2007) and Reed (2010) are two examples of such approaches.

[54] I am in effect endorsing Dana Nelkin's suggestion that it would be a mistake to give disconnected solutions to the

The combination of views that JPK *does* seem to rule out is that justification is possible in lottery cases, but knowledge is not. This is not a completely trivial consequence—some epistemologists have accepted both views. David Lewis seems to have held them.[55] Aidan McGlynn (2012) endorses them too; but McGlynn doesn't consider the combination to be obvious, or a reason to endorse a view that predicts it; his attitude is rather that it is an acceptable consequence of an otherwise-attractive view. I haven't come across many epistemologists who hold it to be obvious that one can have justification but can't have knowledge in such cases. (Should Gettier (1963) have included lottery cases alongside his false-lemmas cases?) So I don't think that JPK has problematic consequences about lottery beliefs.

4.12 History

Another choice point for the JPK theorist concerns the significance of a subject's past history for justification. Consider for example the 'problem of forgotten evidence,' deriving from Pappas (1980) and Harman (1986), and pressed here by Alvin Goldman against internalist theories of justification:

> Many justified beliefs are ones for which an agent once had adequate evidence that she subsequently forgot. At the time of epistemic appraisal, she no longer possesses adequate evidence that is retrievable from memory. Last year, Sally read a story about the health benefits of broccoli in the "Science" section of the *New York Times*. She then justifiably formed a belief in broccoli's beneficial effects. She still retains this belief but no longer recalls her original evidential source (and has never encountered either corroborating or undermining sources). Nonetheless, her broccoli belief is still justified, and, if true, qualifies as a case of knowledge. Presumably, this is because her past acquisition of the belief was epistemically proper. But past acquisition is irrelevant by the lights of internalism.... All past events are "external" and therefore irrelevant according to internalism. (Goldman, 1999, p. 280)

The problem is that facts about the subject's past can be relevant for justification—had Sally read the story in the *National Enquirer* and then forgotten the source, she would now be unjustified—so internalism must be mistaken. The parallel argument against JPK would complain that past facts are relevant for justification, but are not part of a subject's cognitive home. If so, one's belief might be unjustified, because based improperly on past and forgotten evidence, but one might nevertheless be internally alike to a possible person of the relevant sort who has knowledge, based *properly* on past and forgotten evidence.

Whether the objection tells against internalism in general, or JPK in particular, depends on just how one understands internalism or cognitive homes. One common internalist response is to accept the implication and deny Goldman's intuition that the forgotten facts really are epistemically relevant for the purpose of justification.[56] Another possibility, although one I haven't seen much explored in the literature, would be to incorporate past forgotten evidence into the 'internal'—internalism might be understood for instance as a claim of supervenience of justification on *past and present* nonfactive mental states.[57] (A tedious argument about whether this sort of view would count as 'internalism'

knowledge version of the puzzle and the rational belief version of the puzzle; see Nelkin (2000). See also the parallel discussion of Alexander Bird's similar view in Bird (2007, p. 101).

[55] Lewis (1996, p. 551).

[56] Steup (2001, p. 140) and Feldman and Conee (2001, pp. 9–10) take this tack, arguing that cases of true beliefs formed for bad but forgotten reasons are Gettier cases.

[57] Feldman (2005, p. 349) observes that this is a possibility, but does not endorse it. Huemer (1999, p. 352) defends this kind of view, but does not characterise it as a kind of 'internalism'. Moon (2012) argues that this form of 'historical internalism' is susceptible to New Evil Demon challenges. See also Reynolds (2013, p. 374).

would presumably be the result. As I have indicated, I am uninterested in terminological debates about just what it would take to be an 'internalist' view.) Either kind of option is also possible for the defender of JPK.

If the first, more orthodox, internalist response is correct, then beliefs based on bad but forgotten evidence can be justified; the JPK theorist can allow this by pointing to a case of knowledge consistent with the subject's cognitive home; this presumably would be a case that is alike with respect to the present, but where the belief was better-formed in the past. This approach fits well with a narrower conception of one's cognitive home, where long-forgotten facts aren't a part of it.

But there is also a less internalist implementation, corresponding to the broadening of the cognitive home. One can allow that basic evidence and cognitive processing that occurred in the past remain a part of one's cognitive home. A JPK theorist who goes this way will require, for a subject S to have a justified belief, there to be a possible subject S′, alike to S not only in relevant present respects, but also in past ones, who has knowledge. Suppose that some time ago, Rose and Margaret each formed the belief that *p*—Rose, from a very reliable source, forming her belief in a responsible way, and Margaret, by unreasonably trusting an obviously unreliable source. Today, each retains her belief that *p*, but retains no memory about from where her belief derives. Rose's past activity was epistemically appropriate, and Margaret's was not. A JPK theorist who wanted to accommodate Goldman's intuitions can say that (whether or not *p* is true) Rose's current belief is justified, and Margaret's isn't, because there is a possible person, alike to Rose in all relevant respects—*including the virtuous cognitive processing she she displayed in the past*—who has knowledge, while Margaret does not. Margaret can't have such a possible twin because her irrational past cognitive processing—a part of her cognitive home—is inconsistent with having come to knowledge.

4.13 Reliability

Another point of flexibility for the JPK theorist concerns judgments about cases internalists have used to argue that justification and reliability come apart. At first blush, it may appear as if JPK is a competitor to reliabilism in epistemology, as the latter is typically characterised as a theory of justification.[58] However, it is worth remembering that justification is not the only access point into epistemology for reliabilism, and various views that are rightly characterised as reliabilist theories of *knowledge* have also been proposed—see Goldman (2011, section 1) for a catalogue. (Some of these, e.g. that of Dretske (1981), were explicit in denying that the reliabilist component was meant to have to do with justification.) The present view about justification is consistent with reliabilist approaches to knowledge.

Consequently, although JPK certainly fits well with the orthodox position that reliably produced belief is in general insufficient for justification—the position according to which, for example, Lehrer's 'Mr. Truetemp' or Bonjour's clairvoyant Norman do not have justified beliefs, this is not mandatory for a JPK theorist who wishes also to endorse a reliabilist approach to knowledge. If one wants to bite the bullet for reliabilism and say that cases like these can be cases of knowledge and are typically cases of justified belief, there is nothing in JPK preventing one from doing so. This would follow from certain kinds of very close connections between knowledge and reliability. I see no particular reason to think there are such connections, especially once one embraces a knowledge-first stance—but one could certainly develop JPK in that direction if one wished.

Even certain reliabilist approaches to *justification* may be consistent with some versions of JPK. For example, Goldman (1979, p. 17) considers a version of reliabilism about justification according to which a belief is justified just in case it is produced by processes that are reliable in the *actual*

[58] E.g. by Goldman (1979).

world (not necessarily the subject's world). This version of reliabilism entails the supervenience of justification on a subject's cognitive processes; so it is at least as internalist as any version of JPK.[59] There is no obvious conflict between them.

So on the whole, JPK is consistent with a variety of diverse epistemic positions—in particular, it is consistent with all approaches to knowledge.[60] Consequently, it should be acceptable to epistemologists with a wide variety of commitments, so long as they do not hold out hope of reducing knowledge partly in terms of justification. Combined with what I've suggested above—that JPK performs very well with respect to many of the desiderata for a theory of justification, including its intuitive extension—this suggests that there is a strong case to be made in its favour. It deserves serious consideration as a theory of justification.

[59] On the version of reliabilism here considered, someone who uses a process that is reliable in the actual world but unreliable in the world where it is used—a brain in a vat who takes subjective perceptual experience as a guide to external reality, for example—has *justified* beliefs, because they use methods that are *actually* reliable. This is not the standard version of reliabilism.

[60] Very implausible approaches to knowledge, combined with JPK, will result in very implausible approaches to justification too. For example, if knowledge is just true belief, then by JPK all (or at least nearly all) beliefs are justified.

Appendix: Impossible Knowledge, Content Externalism, and JPK

Knowledge requires truth; nothing false is known. Consequently, nothing *necessarily* false is even *possibly* known. A potential worry for JPK, then, is that it may have the implication that necessarily false beliefs cannot justifiably be believed. There is no possible subject who knows that Hesperus is not Phosphorus; therefore, according to the worry, JPK implies that no one can justifiably believe that Hesperus is not Phosphorus. This would be a counterintuitive consequence. The discovery that Hesperus is identical to Phosphorus is attributed to Pythagoras; various of his astronomical observations led him to this conclusion. But it doesn't seem at all out of the question that prior to this discovery, at least many people *were* justified in believing the necessary falsehood that Hesperus and Phospherus were non-identical. And certainly it is not difficult to imagine Pythagoras having made different, misleading observations instead—suppose he'd been the victim of a skilful prank—which would intuitively have justified even him in believing that necessarily false proposition.

I do not believe that JPK is committed to the impossibility of justified belief in necessary falsehoods. Because making this case requires engagement with some controversial and technical issues in the philosophy of mind, I've put the bulk of the discussion of this discussion into this Appendix. Some readers may wish to proceed directly to Ch. 5.

To start with, it should perhaps be pointed out that there is precedent for accepting counterintuitive consequences concerning mental states towards necessary propositions. Stalnaker (1981b) says everyone believes every necessary truth; Lewis (1996) says that everyone *knows* every necessary truth. Approaches linking knowledge to modal properties like safety and sensitivity also, notoriously face challenges regarding knowledge of necessities.[1] Such views are not dismissed out of hand for these features, and the many attempts to explain away the relevant intuitions may apply as well to JPK as to these views. Even still, it cannot be denied that these are counterintuitive consequences. (I argued that Lewis's approach to knowledge should be amended, partly for this reason, in §1.7.) Does JPK have similarly counterintuitive implications? It would if S's justified belief required there to be a possible similar individual who has knowledge *of the very same content* as S's belief. But this requirement is *not* included in JPK. Here, again, is JPK:

> **JPK** S's belief is justified iff there is a possible individual, alike to S with respect to all relevant basic evidence and cognitive processing, whose corresponding belief is knowledge.

Suppose that Pythagorus believes, on the basis of strong but misleading evidence, that Hesperus is not Phosphorus. Assuming that we are working with a modest enough epistemic standard, JPK can allow that Pythagorus's belief is 'justified', by pointing to a possible version of Pythagorus, in

[1] See e.g. Roland and Cogburn (2011) and Blome-Tillmann (2015b).

a world with two distinct heavenly bodies, which Pythagorus's twin 'knows' to be distinct. Twin-Pythagorus knows that *twin*-Hesperus is not *twin*-Phosphorus; this knowledgable belief is the one corresponding to Pythagorus's justified belief that Hesperus is not Phosphorus.

Unlike in the case of Bird's approach, making further inferences from this justified but necessarily false belief poses no special problems, given JPK. If Pythagorus, justifiably believing Hesperus and Phosphorus to be distinct, goes on to infer that Phosphorus is not Venus, this belief can be justified in exactly the same way: he has a possible evidential duplicate who knows that twin-Phosphorus is not Venus. Bird's JuJu, by contrast, has the implication that no judgment based on a necessarily false belief can be justified, for reasons explained in the previous section. So externalism about content can play an important role in rendering JPK's particular internalist approach to justification plausible.

Certain versions of JPK, however, will have to tread carefully here; content externalism is a double-edged sword. If one endorses JPK_i, which requires the counterparts to share the relevant *internal* states, things are relatively simple—twin-Pythagorus is an intrinsic duplicate of Pythagorus, and his internal experiences interact with his environment with a way that leads him to knowledge about his world's celestial bodies. But on a more externalist conception like JPK_p, where the contents of factive perceptual states are basic evidence, content externalism will cause trouble for some interpretations of the latter.

Suppose, for instance, that Hannah has interacted with Robin the farmer, and also with Sir Ruthven the evil baronet, and fails to realise that they're actually one and the same person. Under some circumstances, it seems like it should be possible for her belief to the contrary to be justified. But on the assumption that Hannah's direct perceptual evidence includes propositions like *Robin was holding a plow yesterday* and *Ruthven is cracking a whip today*, it will be difficult to tell a story about a possible alternate version of Hannah who ends up knowing that the farmer and the baronet in her environment are distinct. For JPK_p requires the version in question to share all of Hannah's direct perceptual evidence. The worry is that on orthodox Russellian views about direct reference, the proposition that *Robin was holding a plow yesterday* is *identical to* the proposition that *Sir Ruthven was holding a plow yesterday*. Each proposition has the actual individual as a constituent. If so, one might think that only someone perceptually acquainted with *this particular individual*—the one who is actually a baronet disguised as a farmer—can share Hannah's direct evidence. This will prevent the JPK theorist from following the move sketched above in the Hesperus/Phosphorus case. There, I suggested that twin-Pythagorus may know that two *different* celestial bodies to which he attaches the names 'Hesperus' and 'Phosphorus' are distinct.[2] But if all of Hannah's relevant counterparts have seen one individual under two distinct guises, the factivity of knowledge implies that none of them can know those guises to correspond to distinct individuals.

One possible solution to this challenge would be to give up on JPK_p, opting instead for a more traditionally internalist version like JPK_i. But another option is to finesse one's theory of perceptual content; one needn't hold, just because one thinks there is direct perceptual experience of the external world, that full-blooded Russelian propositions are good candidates for the objects of that experience. Several other options are available. For example, one might take a Fregean approach to content, holding that modes of presentations of objects and properties, rather than the objects and properties themselves, are the direct objects of perception.[3] But one needn't necessarily embrace Fregean contents here; as Susanna Siegal helpfully summaries in her *Stanford Encyclopedia* article on perceptual content, versions of the Russellian view also allow for perceptual contents to be individuated less specifically than by the objects they involve:

[2] I shifted the discussion to one about perceptual recognition of individuals because it is less plausible that particular features of astronomical bodies are *directly* perceived.

[3] See for example Burge (1991), Chalmers (2004), and Silins (2011), all cited in Siegal (2015).

> Instead of strongly Russellian contents just described, one could define weaker Russellian contents that involve the very properties things appear to have, but don't involve the very objects that appear to have them. These weaker Russellian contents are said to be *existentially quantified* contents, which would be naturally expressed by saying something of the form 'There is a red cube at location *L*'. Most philosophers who defend the view that experiences have Russellian contents hold that these contents are weaker Russellian contents. (For general discussion, see Chalmers (2004); for discussion of the case of color, see Thau (2002), Shoemaker (1990), Shoemaker (1994), Shoemaker (2006), Holman (2002), Maund (1995), Wright (2003).) One reason for weakening contents in this way is to respect the ideas that two experiences can represent the world as being the same way, even if they are experiences of perceiving numerically different objects, such as twins.[4]

On this kind of weaker Russellian view, Hannah's perceptual content does not require her to be interacting with this particular individual who is both a farmer and a baronet; as a result, it can be held in common with a possible version of Hannah who is interacting with two distinct people and knows it. And so even JPK_p, suitably developed, is consistent with necessarily false, but justified, belief in this kind of case. Obviously the issues involving the correct theory of perceptual content are vast, and exploring them thoroughly would be a significant sidetrack for the project of this book on knowledge. But it at least appears as if there are versions of JPK_p that are able to avoid the problematic result.

The strategy just sketched shows one way JPK is consistent with the possibility of justified beliefs in necessary falsehoods; beliefs in necessary falsehoods can have true counterparts in possible individuals with the same basic evidence and cognitive processing. How widely does this strategy generalise?

It is straightforward to apply the approach to allow for justified beliefs in other necessarily false identity claims (Clark is not Superman) and natural kind essence claims (water contains no hydrogen). What these saliently have in common is that they constitute the canonical Kripkean necessary *a posteriori*. But can JPK allow for justified belief in *a priori* necessarily false contents? Again, I don't think it's disastrous for JPK if it doesn't. This would be a counterintuitive result, but one one might learn to love. Still, one may not need to. JPK can deliver possible justified beliefs in *a priori* necessary falsehoods if externalism about content extends to *a priori* contents. In particular, it can do so if it is possible for X and Y to share the basic evidence and cognitive processing, with counterpart beliefs B(x) and B(y), where x is *a priori* false, but Y knows y. Standard arguments for social anti-individualism about content do support this possibility.

Consider Tyler Burge's (1979) case of Oscar and his arthritis. Oscar has what Burge calls 'partial understanding' of the concept ARTHRITIS. His understanding is sufficient for us properly to attribute to him beliefs about arthritis, but not sufficient for him to rule out from the armchair that he has arthritis in his thigh. A twin of Oscar, living in a society that categorised ailments differently, might know that it is possible to have twin-arthritis in his thigh, or even that he actually had twin-arthritis in his thigh. So Oscar's belief that he might, or does, have arthritis in his thigh could well be the justified counterpart of this possible knowledge. This even though it is plausibly *a priori* that one cannot have arthritis in one's thigh. (If it were *a posteriori*, it wouldn't be at all clear why Oscar's misapprehension would be require him to have at best partial understanding.)

So JPK allows that subjects with partial understanding, in Burge's sense, can have justified beliefs in a priori necessarily false contents. Only if one's basic evidence and cognitive processing were sufficient to *settle* a given content would justified belief in an *a priori* falsehood have to be counted

[4] Siegal (2015, §3.1)

impossible. However, it is by no means obvious that there are any such cases. Nor is it obvious that there aren't—we are in controversial territory here. So the strategy to follow is certainly intended as a speculative one.

The point is most easily demonstrated within the framework of JPK_i, so let's focus on that version of the view first, even though it's not the one I ultimately prefer. Kripke's (1982) interpretation of Wittgenstein suggests that any subject's intrinsic states won't even guarantee that an arithmetic belief has the particular content it does; if so, JPK_i can even allow for one to have justified false arithmetic beliefs. Suppose that Lionel defers to Conrad's testimony on arithmetical matters, and that in general, this is not an unreasonable idea. (Conrad is good at arithmetic, and trustworthy about these matters.) But suppose that on a particular instance, Conrad makes a mistake, telling Lionel that 57 x 23 = 1461; the unfortunate but predictable consequence is that Lionel believes the a priori necessarily false proposition that *57 x 23 = 1461*. How dreadful! But could his belief nevertheless be justified? It seems to me intuitive that it could (or at least that the justification ascription could be true in contexts where testimony can suffice for satisfying knowledge ascriptions). Although I'd give up the result if that's what it took to retain the otherwise attractive JPK, I'd consider it a mark in its favour if it could extend the story about *a posteriori* necessary falsehoods to these.

If the Kripkenstein idea is right, then JPK_i, at least, can so extend the story. Lionel is the intrinsic duplicate of a possible individual, Lionel′, who has knowledge—not of *57 x 23 = 1461*, for no one could know *that*—but of some other content. Perhaps, following the example of Kripke's 'quus', the content Lionel′ knows is that *fifty-seven quimes twenty-three is one thousand, four hundred sixty-one*, where 'x quimes y' is defined as the product of x and y for all numbers except 57 and 23, in which case it is 1461. Since there is nothing in Lionel's intrinsic state that ensures that he means *times* rather than *quimes*, he has a possible intrinsic duplicate—one, we may suppose, in which his broader mathematical community has done enough to settle the question—who expresses *quimes* thoughts with the beliefs corresponding to Lionel's *times* thoughts.

So I do think, as I argued in Ichikawa (2014, p. 196), that JPK_i can yield justified false arithmetical beliefs. Things may be a bit more complicated for the more general version of JPK, that does not assume that the intrinsic determines the subject's basic evidence. One reason things can be complicated is that it is not clear what one should think constitutes basic evidence in the realm of arithmetic; so for many arithmetical beliefs, it's not at all clear what kinds of possible individuals one should be thinking of. Furthermore, in this *particular* case, the belief is based on testimony. And it is also not clear what one should think is the basic evidence in cases of knowledge via testimony. Reductivist approaches to testimony have it that the basic evidence is perceptual—perhaps Lionel's basic evidence concerns what Conrad said. (And does he get that he asserted the particular *content*, or only that he uttered a certain *sentence*, from which Lionel can infer that he asserted the content?) These questions about the scope of direct perceptual content may end up making differences here, as they did in the discussion of the Robin/Ruthven case in the previous section; as before, a thorough investigation would take us too far afield.

One might well worry (as two OUP referees did) about a theory of justification that commits one to controversial approaches to the philosophy of perception or the philosophy of mind. Shouldn't an epistemological stance try to remain neutral on such questions if possible? I offer two responses. First, as I mentioned above, challenges having to do with necessary contents are common to a wide variety of epistemic approaches, and there are diverse attempts to resolve them in the literature. If one doesn't like mine, others (including ones that bite the bullet and accept that one can't have justified necessarily false beliefs) are also available. But second, to speak more directly to the concern, I want to resist the implicit preference for an insular approach to epistemology that keeps its commitments quarantined from other fields, such as the philosophy of mind and perception. One of the upshots of the knowledge first stance is that theorising about knowledge *will* impact broader questions about the mental. Insofar as my epistemological approach carries commitments in the

philosophy of mind, this is a signal that its interest extends beyond epistemology. At any rate, I see no *obvious* reason why a development of JPK like JPK_p couldn't also tell a kind of a story along the lines given above.

To summarise: since the class of cases in which content supervenes on the relevant base is at best extremely small, and possibly nonexistent, saying that in such cases, justified a priori false belief is impossible is not a significant intuitive cost.

Chapter 5

Action

According to a popular idea, knowledge has important normative roles to play in guiding action. Knowledge, it is suggested, is a (or *the*) 'norm of practical reasoning'. We shall examine more precise versions of this claim below. Some epistemologists have argued to the contrary that reflection on cases suggests that knowledge does not play such roles; others have argued that positing such roles roles stands in tension with contextualism. I disagree with both lines of thought. In this chapter, I have three aims: first, to defend substantive knowledge–action links from opposing arguments in the literature; second, to clarify the best way in which to understand knowledge–action links; and third, to demonstrate that there is no reason that a contextualist faces special problems from knowledge–action links.

As in my dialectical strategy elsewhere in this book, I do not consider it among my primary aims to offer original, positive support for knowledge–action links. Although I will quibble with some of the details below, I think that a strong *prima facie* case for such links has already been given, for example by Hawthorne and Stanley (2008), Weatherson (2012), and Fantl and McGrath (2009). However, it may be helpful to begin briefly with some of the motivations in favour of such links.

5.1 Use of 'Knows'

A first motivation for knowledge–action norms, I mention only to set aside. This is the observation that ordinary assessment of action often proceeds via the invocation of knowledge language. Many defenders of knowledge–action norms have cited this kind of language as providing motivation. For example, John Hawthorne and Jason Stanley write:

> [O]ur ordinary folk appraisals of the behaviour of others suggest that the concept of knowledge is intimately intertwined with the rationality of action. Suppose, for example, that Hannah and Sarah are trying to find a restaurant, at which they have time-limited reservations. Instead of asking someone for directions, Hannah goes on her hunch that the restaurant is down a street on the left. After walking for some amount of time, it becomes quite clear that they went down the wrong street. A natural way for Sarah to point out that Hannah made the wrong decision is to say, "You shouldn't have gone down this street, since you didn't know that the restaurant was here." Hawthorne and Stanley (2008, p. 571)

I agree with Hawthorne and Stanley that this kind of way of criticizing someone is a natural one, but I do not think this counts much, if at all, in favour of substantive knowledge–action links.[1] One

[1] So I agree with McGlynn (2014, pp. 133–4).

reason this is so is that there are many natural ways of criticising action in such cases; Sarah could have said any of the following:

- You shouldn't have gone down this street, since the restaurant isn't here.
- You shouldn't have gone down this street, since you were wrong to think the restaurant was here.
- You shouldn't have gone down this street, since we don't have time to explore based on guesses.
- You shouldn't have gone down this street, since you weren't sure where the restaurant was.

Ordinary folk appraisals do not, it seems to me, support the idea that knowledge has a *special* role to play in justifying actions. Furthermore, as I will argue at length throughout this chapter, intuitions about the propriety of actions and the presence or absence of knowledge aren't typically specific enough enough to have substantive implications about the knowledge norm.[2] So I do not propose to rest argumentative weight on the fact that people are likely to use the word 'knows' when assessing one another's actions. Theoretical principles like knowledge norms must be adjudicated on theoretical grounds.

5.2 Reasons

Another motivation for knowledge–action links begins with the simple observation that there is a distinction between the abstract *existence* of some reason for someone to act, and someone's *having* a reason to act. This distinction is sometimes described as a difference between 'objective' and 'subjective' reasons, as by Mark Schroeder in his (2008) discussion of this famous example from Williams (1979, p. 60):

> [Bernie] is standing around at a cocktail party, holding a glass that he believes to contain the gin and tonic for which he has just asked his hostess, but which in fact is full of gasoline. There is a reason for Bernie to set his glass down without taking a sip—namely, that it is full of gasoline. What better reason to avoid taking a sip could there be? But this is not a reason that Bernie *has*, in the sense we've articulated, for he is unaware of it...

Schroeder calls reasons a subject might be unaware of—like the fact that the glass contains gasoline—'objective reasons'. He calls the kind of reason that one has to have in mind 'subjective reasons'. Although the distinction described here is surely a real one, I do not prefer the labels 'objective' and 'subjective' to describe it; 'subjective' runs the risk of suggesting that these reasons are purely a matter of the subject's beliefs or perspective. But it is possible to be radically mistaken about which reasons one has, even in the so-called 'subjective' sense. Similarly, calling something an 'objective' reason may implicate that it obtains in a mind-independent way, and that its status does not depend on any of the subject's contingencies.[3]

[2] Locke (2014, p. 85) emphasises that some of the arguments I have given in defence of the knowledge norm also undercut certain possible motivations for it. I agree with Locke about this. (Indeed I said as much in Ichikawa (2012, p. 54)). Although I embrace knowledge norms, looking at cases like these is not a good way to argue for them.

[3] As e.g. in Jennifer Carr (2015), who writes: 'If you think the apple is poisoned, there's a sense in which you *ought* not eat the apple—even if, unbeknownst to you, the apple isn't poisoned. This *ought*, the subjective *ought*, isn't just sensitive to sources of value in the world: it's also sensitive to what information is available. The objective *ought*, by contrast,

Unlike Schroeder, rather than trying to mark a distinction between *kinds* of reasons, I prefer to talk of the distinction between there *being* a reason for some action, and one's *having* such a reason. (The former may correspond to so-called 'objective' reason, and the latter to 'subjective' reason.)[4] Once the question is framed in this way, it appears very plausible that *having* a reason is an epistemic matter. What kind of epistemic access to a consideration is required in order to 'have' it as a reason? Knowledge is a natural candidate.[5] It is not the only possible candidate—one might hold that to have p as a reason, one need only believe p, or believe p with truth and/or justification.[6] One might think that one's reasons are all and only that which one is justified in believing that one knows.[7] Or one might demand something stronger than knowledge—perhaps one must know that one knows p, or one must be *completely certain* that p. All of these are possible competitor views; however, none of them seem obviously to perform better with respect to the intuitive extension of a subject's possessed reasons. And arguably, the hypothesis that knowledge is the relation in question is more natural than these competitors, if the broad thought of the knowledge first program is correct.[8]

The dialectic here is straightforwardly similar to that involved in the discussion of evidence and knowledge. This should not be surprising; the notions of evidence and reasons are deeply connected. At a first approximation, one might identify one's evidence for p with the reasons one has to believe p.[9] Although I will not pursue the relationship between reason and evidence further here, the idea that reason and evidence should each receive a similar treatment does seem theoretically parsimonious. (I'll return to the relationship in more detail in §5.11.)

The version of the knowledge–action principle I will be examining, and defending, in this chapter, is given in terms of reasons. Here it is:

> KR: Subject S has proposition r as a reason for action if and only if S knows r.[10]

One sometimes sees the knowledge–action principle articulated in terms of what is *appropriately treated* as a reason for action:

---

is insensitive to knowledge and ignorance: it's the *ought* from a God's-eye view.' I don't think that Carr is dividing the questions up in the right way; the 'ought' that is related to the kinds of reasons at issue in this chapter is not insensitive to knowledge and ignorance; but it is also not settled by belief. Instead, it is a matter of *knowledge*.

[4]Schroeder seems to think that views like this amount to 'factorizing' approaches to having reasons, but I disagree. I am *not* suggesting that the fact that S has r as a reason is *grounded in* the independent existence of reason r, combined with S's access to r; it is very plausible that the order of explanation runs the other way round: S's having r as a reason grounds the existence of reason r. While (as I will discuss below) I disagree with Schroeder that there are two independent kinds of reasons, I do agree that it would be a mistake to factorise having a reason into the independent existence of a reason that is, moreover, 'had'. In this respect, having reasons is, to use Schroeder's example, more like having golf partners than it is like having opera tickets. See also Hawthorne and Magidor (2016, §3) for a similar treatment of Schroeder's case, although note that they seem to be using a weaker notion of 'factorizing' that is neutral on the grounding question.

[5]Hawthorne and Magidor (2016) defend this version of the view.

[6]Schroeder (2011) suggests it's even weaker than belief. He holds that any 'presentational attitude'—including belief, but also including non-veridical perceptual experience—suffices for reason-having.

[7]Neta (2009) defends a view in this neighbourhood.

[8]I see an affinity between my approach here and that in Brian Weatherson's (2012) 'Knowledge, Bets, and Interests'. In that paper, Weatherson observes that when representing choices in decision theory or game theory, one must be relying on some kind of epistemic assumption, and he suggests that knowledge is the best candidate. For example, one should include in a matrix representing a subject's decision all and only the payoffs that the subject *knows* to obtain, and the choices the subject *knows* to be live. There is also an affinity with some early work by John Hyman; in his (1999), he suggested that we should understand knowledge as an ability to do things for factual reasons.

[9]One reason this is only an approximation is that there might be non-epistemic reasons to hold a belief—for example, if one is offered some kind of reward for believing that p.

[10]McGlynn (2014, pp. 59–60) calls something very close to this principle the 'Unger–Hyman' thesis, citing Unger (1975, p. 200) and Hyman (1999).

> KATR: Subject S can appropriately treat proposition *r* as a reason for action if and only if S knows *r*.[11]

On the assumption that one can appropriately treat a proposition as a reason if and only if one has that proposition as a reason, KATR and KR are equivalent, and KR seems the better candidate for being more fundamental. (That it is among your reasons is the explanation why it's appropriate to treat it as such.) On the other hand, if it can be in the relevant sense 'appropriate for one to treat *p* as a reason' even though *p* is not a reason one has, KR and KATR will come apart; in such cases, KR seems preferred.

One also sees knowledge norms articulated by reference to 'premises in practical reasoning':

> KPPR: Subject S can appropriately use *r* as a premise in practical reasoning if and only if S knows *r*.[12]

I take it that relying on something as a premise in practical reasoning just is treating it as a reason. So I will treat KPPR and KATR as equivalent.

Another way of drawing knowledge–action links—that articulated by Fantl and McGrath (2009)—is importantly different, for reasons that will come out in §5.5 below.

5.3 Contextualism and Norms

Contextualism is sometimes thought to stand in tension with normative roles for knowledge. Here is a challenge similar to ones pressed by John Hawthorne and Timothy Williamson. Take a case where a subject satisfies low knowledge standards, but not high ones. For example, suppose that Strephon's mother introduces him to Celia as his aunt. In fact, she is his aunt, and no one is engaging in any epistemological shenanigans. Relative to many ordinary conversational contexts, sentence (1) is true.

> (1) Strephon knows that Celia is his Aunt.

But on the contextualist approach I have been defending, there will be skeptical conversational contexts where (1) is false. For example, in a context in which it is treated as an open question whether Strephon's mother is lying to him, assuming he has no evidence—or more carefully, in light of ch. 3, nothing satisfying 'evidence' in that context—ruling that possibility out, (1) will express a falsehood, and (2) instead will express a truth:

> (2) Strephon does not know that Celia is his Aunt.

For concreteness, suppose that Mountararat is in a low-standards context and utters (1), and Tolloller is in a high-standards context and utters (2). As outlined in ch. 1, one of the standard virtues of contextualism is its ability to respect the intuitiveness of both utterances, dissolving the appearance of disagreement. However, the objection goes, the contextualist approach resolves *too much* disagreement. Given knowledge–action links, the contextualist approach threatens to dissolve apparently *substantive* disagreements about what actions are appropriate.

But we must bear in line the lessons of the previous chapters; we must take care not to make use–mention errors. In particular, we must *not* argue thus: 'knows' is context sensitive; one's reasons are all and only one's knowledge; therefore, one's reasons are context-sensitive. To do so would be

[11] Cf. Hawthorne and Stanley (2008, p. 578).

[12] Cf. Williamson (2005a, p. 231) Hawthorne (2004, p. 30).

to conflate semantics with metaphysics. If knowledge-action links and contextualism about 'knows' lead to pressure for contextualism about normative language, the connection isn't a direct one. Here are two ways of attempting to set out this kind of pressure. One is to observe that people who attribute knowledge that *p* will also tend to think it appropriate for one to act as if *p*. For example, suppose that after Strephon is told that Celia is his aunt, he responds by embracing her. (Suppose also that the social context is such that her being his aunt would be, if possessed, a good reason to embrace her.) Mountararat, who is in a low-standards context and uttered (1), will presumably take Strephon's action to be perfectly justified. By contrast, Tolloller, who is in a skeptical context, unwilling to take it for granted that no one is trying to deceive Strephon, might well think that Strephon is being a bit hasty, not merely in accepting the testimony, but in embracing Celia. They may express their judgments by uttering (3) and (4), respectively:

(3) Strephon was right to embrace Celia.

(4) Strephon ought not to have embraced Celia.

On the face of it, (3) and (4) are in conflict; it seems like Mountararat and Tolloller are having a substantive disagreement about what Strephon ought to do. However, the objection continues, it is not plausible to apply the contextualist diagnosis to *this* disagreement.[13]

Why is it implausible to apply the contextualist diagnosis to the (3)–(4) disagreement? 'Ought' is, after all, a context-sensitive modal; it's clear that there is *possible* merely apparent disagreement in cases like this.[14] One may suppose, for example, that in some contexts, very scrupulous standards are operative—*by the lights of the law*, one ought to come to a complete stop at a stop sign, whether or not there is traffic around, even if *by the standards of safe driving*, slowing down enough to make sure there's no oncoming traffic or pedestrians may be just fine. Carrie Jenkins and Daniel Nolan, for example, have argued that apparent moral disagreements between maximizing and satisficing contextualists ought to be resolved in this way.[15] While it's clear that 'ought' language is context-sensitive, it's not clear to me that the objection we're considering is so easily diffused. The Jenkins/Nolan move, it seems to me, is applicable when we're trying to consider, of some sub-optimal action, whether it's sufficiently bad to warrant the normative sanction. But in the case of Strephon, the (apparent) disagreement between Mountararat and Tolloller isn't about whether to sanction some action that is uncontroversially *to some degree* deficient; it's about whether there's any normative failure at all. Mountararat, in saying (1) and (3), isn't merely saying that Strephon is *close enough* to perfection to justify positive normative predicates; he's saying that he actually did everything just right. So I agree with the critics of contextualism who think that a kind of context-sensitivity about the normative language is not the answer to this puzzle.

This version of the argument turned on the prediction that the kind of person who would want to consider (1) true would also consider (3) true, and that someone accepting (2) will tend also to accept (4). Whether someone who accepts (1) really will also accept (3) is a matter of that individual's psychology; another way of developing a similar sort of objection argues that it's problematic to separate them. The following version, which is closer to the form of the argument that John Hawthorne gives, involves the attribution of statements of knowledge-action links to speakers assessing someone's actions.

[13]The worry here is directly inspired by the challenge Williamson presses for contextualists in Williamson (2005b), although I have adapted it for the case of action. (Williamson's example involves a disagreement about whether one ought to believe, given the knowledge norm of belief.) A similar issue is pressed in Hawthorne (2004, p. 88). (Hawthorne *is* writing about action norms.) The challenge is quite similar to one advanced against contextualism about deontic modals; see e.g. Kolodny and MacFarlane (2010). But see also Bronfman and Dowell (2016).

[14]See e.g. Bronfman and Dowell (2016).

[15] (Jenkins and Nolan, 2010). See also Unger (1995).

Suppose as before that Tolloller is in a skeptical context, and so is inclined to treat (2) as true. However, suppose that he *doesn't* go on to criticise Strephon's embrace; he accepts (3) instead of (4). Perhaps he recognises that Strephon himself is not in a practical or conversational situation where it is important for him to consider the possibility that his mother is lying, even though Tolloller himself is operating with higher standards.[16] Certainly this is a possible psychological state in which one might be. The problem, according to Hawthorne, is that within his high-standards context, Tolloller will have to conjoin his verdicts in ways that are at odds with statements of theoretical principles to which he might want to be committed. For example, if, relative to his context, (2) is true, but Strephon was nevertheless perfectly reasonable in acting as he did—embracing Celia on the grounds that she was his Aunt—then it seems that this should be true in Tolloller's context:

> (5) Strephon doesn't know that Celia is his Aunt, but he is perfectly reasonable in embracing her for the reason that she's his Aunt.[17]

Hawthorne calls something like this implication 'surprising, to say the least' (p. 88), and describes such cases as examples of 'sentences that seem intuitively incorrect' (p. 89). Myself, I do not place much stock in very theoretical intuitions like this one. Whether certain highly strained, theoretical English sentences sound odd, abstracted from complicated conversational contexts, doesn't seem like something epistemologists should be too worried about. It's a Moorean fact that many ordinary knowledge ascriptions are true, and it's a theoretically attractive idea knowledge is systematically related to rational action. But that sentence (5) should be false in Tolloller's context seems to me neither.[18] I just have no intuitive verdict about (5). I don't consider having to accept that it is true in certain contexts when describing certain situations to be 'surely a cost of the view'. But I shall return later in this chapter to the question of whether contextualists are committed to its truth in high-standards contexts. Before doing so, however, I'd like to set aside challenges specific to contextualism, and examine some objections to the knowledge norm itself. It is useful to take a bit of a detour here because it will turn out that the best response to some of these challenges for any defender of KR (whether or not a contextualist) has a good deal in common with the contextualist response to the relationship between KR, contextualism, and contextualism about normative language.

There are two kinds of apparent counterexamples to KR: counterexamples to the *necessity* of knowledge for having a reason, and counterexamples to the *sufficiency* of knowledge for having a reason. I'll consider these classes, in the next two sections, in turn.

5.4 Intuitive Counterexamples to Necessity

The main kind of case that is sometimes thought to count against the necessity of knowledge for reason has these features: a subject has a belief that does not amount to knowledge, but where its failure to be knowledge is in some way hidden from the subject. The subject performs the action

[16] On certain ways of spelling out this kind of case, the knowledge norm of assertion or belief will complicate the story—if Tolloller doesn't know by his high standards that Strephon even knows by low standards (perhaps because he doesn't know by his high standards that Celia *is* his aunt), then he might not be willing to outright assert that Strephon acted reasonably—maybe he should only say that he might have done so. I think the cleanest way to get around this complication is to suppose, if necessary, that Tolloller has more evidence than Strephon does, so that *he* knows even by his own high standards.

[17] Hawthorne's own version of this consequence is the more schematic "you should rely on propositions that you don't know to be true in your practical reasoning". (Hawthorne, 2004, p. 88) For why putting it this is problematic, see DeRose (2009, pp. 253–4).

[18] There is a point of affinity here with Keith DeRose's 'methodology of the straightforward', according to which which we should be much more confident in our intuitive verdicts about simple cases than about highly theoretical notions like those at play here. DeRose (2005, pp. 153–4).

that would be made reasonable if the belief really were knowledge; intuitively, the action really is reasonable, so (the argument goes) belief contents can be possessed as reasons even when they aren't known.

For example, suppose that Phyllis has strong but misleading evidence that suggests that her fiancé, Strephon, has been unfaithful to her. As a result, she decides to break off her engagement, and accept the suit of a rival. Phyllis believes that Strephon has been unfaithful, but she doesn't know it, for he hasn't. (There is a perfectly innocent explanation for the compromising position in which Phyllis found him, albeit one that seems rather incredible from her point of view.) The thought is, it is reasonable for her to break off her engagement, even though her purported reason—that Strephon has been unfaithful—isn't known. It seems that one must either admit that reason possession does not require knowledge, or deny that Phyllis acts reasonably under the circumstances. Given the preponderance of cases like Phyllis's, this latter option would seem to require a rather extreme form of externalism about reasonable action: Phyllis's possible counterpart—who really does know that her fiancé has been unfaithful—may reasonably break off her engagement, but Phyllis herself would be unreasonable to do so. This, according to the objection, is too big a pill to swallow.[19](Just which possible versions of Phyllis are in the relevant sense counterparts will depend on one's broader commitments about basic epistemic states—see Ch. 3. One possible view will consider possible individuals *intrinsically identical to Phyllis*; another will consider those who have *the same basic evidence and cognitive processing* as Phyllis. This choice point corresponds directly to the various ways of spelling out approaches to evidence and justification in the previous chapters. As in the previous chapter, I'll use 'cognitive home' as a placeholder for whichever features of Phyllis play the fundamental epistemic role in question.)

I agree that this result is unacceptable; I do not agree, however, that KR commits one to such a conclusion. Note that the following argument is invalid:

1. KR implies that what reasons S has is not settled by S's cognitive home.

2. What reasons S has wholly determines what actions it is rational for S to take. Therefore,

3. KR implies that what actions it is rational for S to take is not settled by S's cognitive home.

Premise (1) follows directly from the statement of KR and the assumption that one sometimes has non-trivial knowledge of the non-domestic world. Premise (2) is plausible, and I do not wish to deny it. But the externalist conclusion, (3), does not follow. It is possible for there to be distinct bodies of reason that nevertheless commend the same action. So a difference in reason does not imply a difference in what is reasonable. Although Phyllis does not know that Strephon has been unfaithful, we shouldn't overlook the fact that she *does* have a good deal of relevant knowledge—i.e., whatever evidence she has that makes it reasonable, if mistaken, to believe that he's been unfaithful.[20] Unless there is some reason to suppose that *these* reasons wouldn't rationalise her action, KR faces no pressure to accept (3). (I'll examine some attempts to argue that such reasons wouldn't support such actions below.)

Although this is a relatively simple point, I do think that it is sometimes overlooked. For example, Jessica Brown and Mikkel Gerken have each argued in recent papers that, since someone in a Gettier

[19]For something like this argument, see Gerken (2011), who argues that the knowledge norm commits one to an error theory about this kind of result, and Schroeder (2008), who argues that Phyllis must have the false proposition that Strephon has been unfaithful as a reason.

[20]In a similar move, Hawthorne and Stanley (2008) cite subjects' knowledge of probabilistic contents.

case is intuitively reasonable in performing the same actions as is someone with knowledge, the knowledge norm implies an error theory.[21] For example, Brown writes:

> I criticise my partner for failing to get potatoes on his way home. ... [H]e calls in at the supermarket to get beer but fails to buy any potatoes, despite knowing that we were out. I criticise his action saying: 'You should have got potatoes. You knew we were out'. It seems that this criticism would not be undermined if it turned out that he didn't know we were out of potatoes but was in a Gettier situation in which he merely had the justified true belief that we were out. (Brown, 2008c, p. 173)

Whether the criticism would be undermined depends on what *precisely* the criticism is. If the criticism is the quoted utterance—'You should have got potatoes. You knew we were out'—then it *would* be undermined, for the straightforward reason that it would be false. But Brown is right that there is something significant that is not undermined—her partner isn't wholly absolved by being in a Gettier case. He still acted improperly, and Brown's charge is that KR is unable to explain that fact, since he didn't know they were out of potatoes. But is this the kind of failure that KR speaks to? This is less clear. On at least many plausible ways of spelling out Brown's case, her partner is criticisable, not for failure to infer according to the knowledge norm, but because he is thoughtless, careless, or selfish. For example, consider a different version of the case: suppose that although he knew they were out of potatoes, he didn't buy any more because he cared more about having beer than tackling the shopping list, and didn't want to carry both. In this version, assuming their domestic arrangements give him responsibility for such shopping, he *is* criticisable, but not for violating the knowledge norm. In this version of the case, he never violates the knowledge norm. Instead, he is criticisable for putting his own selfish desires over the needs of his household; KR is consistent with there being many other norms on action than the knowledge norm. So even if KR does not itself directly explain why Brown's partner is being a jerk, this needn't be a problem for KR. Maybe he is being a jerk in a way that has nothing to do with which propositions he is treating as reasons. Although we're generally good at intuitively recognising people who are doing wrong, we're much less good at identifying what exactly is the norm that they are violating.[22] Any attempt to rebut KR in a case-based way has to reckon with this difficulty. The details of the wrongdoing matter a lot.

Even if, making the case as helpful to Brown as possible, we suppose that there is nothing criticisable in his affect, and that his primary desire was to fulfil his domestic responsibilities wishes vis-à-vis the shopping, it is still unclear that the central intuitions Brown relies upon are at odds with KR. Remember once again that subjects of Gettier cases *do* have lots of knowledge. For example, in the version of the Gettier case that Brown goes on to describe, her partner knows that a list, apparently left by Brown to inform his shopping trip, indicates 'potatoes' as to be purchased. (He doesn't know that they're out of potatoes, as the list turns out to be part of a juvenile prank.) But KR does have it that the proposition that potatoes are on the list is among the reasons he has; this, one might well think, is a strong enough reason to justify buying potatoes, or to warrant criticising him for failing to do so.

So it is a mistake to infer from KA's implication that external factors are relevant for what reasons a subject has to the conclusion that external factors alone can make for a difference in what a subject ought to do; the reasons that are absent in skeptical scenarios and Gettier cases may well be

[21] Brown (2008c, pp. 171–2); Gerken (2011, pp. 535–6). In fairness to Brown and Gerken, their most specific targets, Stanley and Hawthorne, do seem to want to embrace such an error theory. See Gerken's pp. 538–9, citing Hawthorne and Stanley (2008, p. 286). If I am right, then Hawthorne & Stanley needn't have committed to an error theory.

[22] For one rather strong defence of this kind of thought, see Haidt (2001). For a retort (that is still consistent with my general observation), see Fine (2006). See also a related thought in Star (2008, p. 333).

redundant in the sense that, even without them, the same actions are permitted and required as they would be if those reasons were present. In §5.6 below I will consider a version of the Brown/Gerken argument that proceeds from stronger premises about for which particular reasons subjects explicitly act; in §§5.9–5.10 below I will consider arguments to the effect that this kind of redundancy does not obtain. But first, I would like to consider intuitive counterexamples to the *sufficiency* of knowledge for reason. As we will see, the dialectic there is quite similar to that here.

5.5 Intuitive Counterexamples to Sufficiency

An apparent counterexample to the sufficiency of knowledge for reason would be a case where a subject intuitively knows that p, but who nevertheless would intuitively not be reasonable in relying on p in her practical reasoning. Remember that 'relying on p in one's practical reasoning' is a theoretical notion; it's difficult to have very clear intuitions about whether this condition obtains. As a proxy, therefore, critics of knowledge norms often focus their attention on establishing cases where a subject knows that p, but where it is nevertheless intuitively unreasonable for the subject to act in a certain way. The difficulty for this strategy is that there is a substantive gap between such cases and counterexamples to KR. This is easily seen via reflection on cases like the following.[23]

> *Fairy Law*
>
> It is illegal for fairies to marry mortals; this crime is punishable by banishment. At one point during her investigation into whether this crime has been committed, the Fairy Queen observes that the defendant is in an unusually good mood. She suspects, but does not know, that she is in a good mood because she married a mortal. At this stage in the investigation, it would be unreasonable of the Fairy Queen to banish the defendant.

There is nothing particularly unusual about the Fairy Queen's epistemic or practical situation here; she knows some things (whether the defendant is in a good mood), but there are other relevant things she doesn't know (whether the defendant is guilty); all told, her epistemic position does not warrant a certain action (banishment). The facts of the case are very ordinary; they do not imply anything at all directly about KR. Notice, however, that the facts of the case include these two:

> (6) The Queen knows that the defendant is in a good mood.
>
> (7) It would be unreasonable for the Queen to banish the defendant without first collecting more evidence.

A clear moral, then, is that pairs of intuitions like these—one an intuition attributing knowledge, and one an intuition denying the rationality of a certain hypothetical action—themselves imply nothing interesting about KR. The obvious reason for this is that KR is only a principle about what reasons one has; it does not by itself say anything about which actions are reasonable. Only in conjunction with a substantive ethical story relating particular reasons to particular actions could KR have implications about what one ought to do. In this case, KR is straightforwardly consistent with these two verdicts given only this very plausible ethical assumption:

[23] I made the same point with a similar example in Ichikawa (2012, p. 49); much of the discussion in this section rehearses and develops ideas from that paper. The point is also related to the strategy of Boyd (2015), who emphasises the difference between satisfying KA and 'acting in an epistemically responsible way'.

(8) That the defendant is in a good mood is not a sufficient reason for the Queen to banish her.

KR says that since she knows it, the Queen has, among her reasons, that the defendant is in a good mood. It is by itself silent about what is a good reason for what. So in general, a case in which one knows that p, but doesn't have enough evidence to Φ, is not an objection to knowledge norms. So much is, I should think, obvious and uncontroversial, when applied to cases like Fairy Law. Nevertheless, I do think that some epistemologists tend to overlook this dialectical situation when presenting putative counterexamples to the sufficiency of knowledge for reason.

Here is one example, given by Jennifer Lackey.[24]

> *Nomination*
>
> Rowena and Cesar have been colleagues in the English Department together for 7 years, and they have a great deal of respect for one another's evaluations of students in the program. While having lunch together last week, they were discussing the progress of various majors in the department. Just prior to dashing off to their respective classes, Cesar told Rowena that Mitchell Jones has excellent writing skills, which Rowena accepted despite the fact that she has had Mitchell in class for only a few weeks and has yet to see any of his writing herself. On the basis of Cesar's reliable and trustworthy testimony—combined with her background knowledge that the possession of excellent writing skills is sufficient for being a candidate for the departmental writing award—Rowena nominated Mitchell for this award in the English Department and thereby agreed to serve as an advocate of this nominee. (Lackey, 2010, pp. 364–5)

Lackey argues that in this case, the following are both true:

(9) Rowena knows that Mitchell has excellent writing skills.

(10) It was unreasonable for Rowena to nominate Mitchell for the award, given her epistemic position.

On these grounds, she concludes that knowledge is insufficient for grounding action.[25] One possible response that the contextualist defender of KR has is to exploit the context-sensitivity of 'knows' and argue that, in contexts where one emphasises the epistemic deficiency that makes nomination unreasonable, Rowena won't count as satisfying 'knows that Mitchell has excellent writing skills'. But using this manoeuvre to respond to such objections will make the Hawthorne/Williamson challenge mentioned above seem that much more trenchant: what should we think about the propriety of Rowena's action, independently of the particular epistemic standards that are operative in a given conversational context? At any rate, no such machinations are necessary; let's continue to let contextualism sit idly by, working in the object language in a context in which (for the purpose of argument we may grant) the statements of Lackey's intuitions express truths. There is still no challenge to KR. For just as in the case of Fairy Law, the two intuitions posed about *Nomination* are themselves *wholly consistent* with the knowledge norm. KR does not predict that the mere observation that Rowena

[24] I discuss this and two other cases in Ichikawa (2012). The others are Brown (2008a, pp. 1144–5) (which I also discuss below) and Reed (2010, p. 232).

[25] My defence of KR leaves open that Lackey successfully shows that some kinds of knowledge—isolated second-hand knowledge, for example, is insufficient to rationalise some actions. I have no quarrel with her claim that proper action in such cases requires more. (Lackey has indicated to me in conversation that this was her central intended claim.) What I want to insist, however, is that in such cases, the content known is genuinely possessed as a reason—if it's insufficient to rationalise the action, it's because that reason isn't a good enough reason, not because it isn't held as a reason.

knows a particular fact implies that she'd be reasonable in taking a particular action; it does so only on the assumption that the fact in question is a sufficient reason to rationalise the action. If this relationship between reason and action isn't established, there is no conflict between these intuitions and the knowledge norm. Now Lackey does say a bit more that is of some relevance here, which might be thought to distinguish *Nomination* from *Fairy Law*: she stipulates that 'Rowena knows that possession of excellent writing skills is sufficient for being a candidate for the departmental writing award'. But this isn't enough to get around the objection I'm pressing—that the following argument form is patently invalid, whether or not one includes the second premise:

1. *r* is among *S*'s reasons.
2. *r* implies that *p*. Therefore,
3. *S* is reasonable in performing action Φ.

This is transparently invalid in general; but it is also invalid in the special case where *p* is that someone is worthy of an award, and Φ is nominating that person for the award. Notice, for example, that it is consistent with all of Lackey's stipulations that only the particular facts in virtue of which a student's writing is excellent would be appropriate reasons to justify making a nomination. Other candidate requirements are available as well; maybe it is only reasonable to volunteer to serve as an advocate for a student if one has *first-hand* knowledge of the student's work. Indeed, Lackey herself seems to think something like this, when she observes that '[i]f the other members of the department asked for an explanation or defence of her action, they would be entitled to feel resentful when she informs them that she cannot offer any further data or support for this because it is grounded entirely in an isolated statement made by one of her colleagues.' (Lackey, 2010, pp. 366–7) One very natural way to encode this fact within the framework of KR, and consistent with Lackey's intuitions, is at the level of which reasons justify which actions: *that he is an excellent writer* just isn't sufficient reason for Rowena to nominate him, even though it is a reason that she genuinely has.

Exactly the same holds for Jessica Brown's widely-discussed example of a surgeon who checks a patient chart, even though she already knows which kidney requires removal.

> *Surgeon*
>
> A student is spending the day shadowing a surgeon. In the morning he observes her in clinic examining patient A who has a diseased left kidney. The decision is taken to remove it that afternoon. Later, the student observes the surgeon in theatre where patient A is lying anaesthetised on the operating table. The operation hasn't started as the surgeon is consulting the patient's notes. The student is puzzled and asks one of the nurses what's going on:
>
> > Student: I don't understand. Why is she looking at the patient's records? She was in clinic with the patient this morning. Doesn't she even know which kidney it is?
> >
> > Nurse: Of course, she knows which kidney it is. But, imagine what it would be like if she removed the wrong kidney. She shouldn't operate before checking the patient's records.[26]

Brown argues that Surgeon is a counterexample to the sufficiency of knowledge for reason, since it is a case where the surgeon knows that the left kidney is diseased, but where it would be unreasonable for her to operate without double-checking first. That is to say, this is another attempt to argue

[26] Brown (2008a, pp. 1144–5) and Brown (2008c, p. 176).

against KR by pointing to cases where S knows that p but ought not to Φ. But we needn't assume that p is a sufficient reason to Φ. KR implies that the proposition that the left kidney is diseased is, if known, among the surgeon's reasons. (As before, we will set aside the distinctively contextualist manoeuvre of arguing that high standards govern discussion of such life-and-death situations, relative to which the attribution of knowledge to the surgeon is false; we also set aside the pragmatic encroachment strategy, which also denies the truth of the knowledge-attributing intuition.) But KR is totally silent on whether that proposition justifies operating in this situation; it is consistent with all of Brown's stipulations that in order for it to be reasonable to remove the left kidney, the surgeon would need stronger reasons. Any of these, for instance, seem to me to be candidates for the kind of reason she'd need to possess to justify beginning operating:

s_1 The chart indicates that the left kidney should be removed.

s_2 The surgeon read the chart.

s_3 The surgeon has signed the box indicating that she has read the chart.

s_4 The surgeon is legally authorised to begin surgery.

s_5 The surgeon is permitted by hospital policy to begin surgery.

s_6 The surgeon knows that the left kidney should be removed.

None of these need be predicted by KR to be held as reasons before the surgeon reads the chart. Certainly s_2 isn't known, because it is false. And unless the surgeon is lying (which presumably would not be a permissible action for independent reasons), s_3 will be false as well, and therefore unknown. It is also reasonably plausible to suppose that the surgeon doesn't know what the chart says until she reads it, in which case s_1 too would be unknown. The other possibilities mentioned here are a bit more complicated. Whether s_4 or s_5 are known, or indeed true, depends on the details of the regulations that apply in the surgeon's environment. But if for example the relevant governing bodies or organisations mandate compliance with typical U.S. health accreditation standards, then s_4 or s_5 will be false, and therefore unknown, until the surgeon checks the chart.[27]

As for s_6, intuitions may well vary as to whether the surgeon knows it prior to checking the chart—although I am going along with Brown's stipulation that she *knows* that it is the left kidney, it doesn't follow from this that she *knows* that she knows it, which is what it would take, according to KR, to possess s_6 as a reason. I also don't have particularly clear intuitions as to whether s_6 would be needed to be a good enough reason to proceed—though note that Williamson (2005a) suggests that iterated knowledge of this sort may be just what is necessary in high-stakes situations like this. My point in including it on the list is to demonstrate the rather onerous burden one takes on by attempting to provide case-based counterexamples to KR. Brown is tacitly assuming that *none* of s_1–s_6 are needed to rationalise the operation; she apparently considers it obvious that the proposition

[27]The protocol in question is the 2004 'Universal protocol for preventing wrong site, wrong procedure, wrong person surgery.' These guidelines specify that a surgeon must 'ensure that all of the relevant documents and studies are available prior to the start of the procedure and that they have been reviewed and are consistent with each other and with the patient's expectations and with the team's understanding of the intended patient, procedure, site, and, as applicable, any implants,' and explicitly call for an 'ongoing process of information gathering and verification, beginning with the determination to do the procedure, continuing through all settings and interventions involved in the preoperative preparation of the patient, up to and including the "time out" just before the start of the procedure.' (Best practices also call for an unambiguous indication with permanent marker on the subject's body.) In a recent article, Mehtsun et al. (2013) identified 2,413 instances of wrong-site surgeries in the United States between 1990 and 2010; the median malpractice payment was $43,197 US. See Rogers et al. (2004) for some discussion of the history and limitations of such norms. Thanks to Sarah Little and Brian Weatherson for drawing my attention to these guidelines and studies.

that the left kidney is diseased would suffice. But this is not at all obvious; it is a deep substantive commitment.[28]

So one of the central moves in my defence of KR against both classes of counterexamples has been to argue that, absent particular commitments from an ethical theory about what reasons would justify what actions, the kinds of cases often proposed are inconclusive.[29] Later in this chapter I will examine some attempts to fill in details of an ethical theory that some critics have thought could bolster the case against KR. First, however, I'd like to consider versions of the counterexample-based argument that attempt stronger stipulations, or that claim stronger intuitions, that may be thought to avoid the need to invoke the details of an ethical theory. So far I have been supposing that the central intuitions concern what is known and whether the actions are appropriate in particular situations. But the dialectic is different if more details are claimed about the cases. Examining some such stronger claims is the subject of the next section.

5.6 More Specific Theoretical Intuitions

In his (2014) criticism of my (2012), Dustin Locke argues that I am mistaken in understanding counterexample-based arguments like those discussed in the previous sections as relying principally on intuitions about the appropriateness of particular actions. Instead, Locke suggests, the counterexample-based arguments proceed directly on the basis of assessments of patterns of *inference*. Discussing Brown's Surgeon case, Locke writes that 'Brown rests her argument directly on the claim that it is not appropriate for the surgeon to premise that the left kidney is diseased in her practical reasoning'.[30] As Locke observes, there is textual support for this interpretation of Brown; as he quotes, Brown does claim that 'the relevant intuition in SURGEON is that the surgeon should not rely on the premise that it is the left kidney which is affected in practical reasoning' Brown (2008a, p. 1145). However, I do not think the dialectic is as simple as Locke supposes it is; although she does make the claim indicated, I read it as her way of explaining or justifying the more direct evaluative intuition, which is about action. This is evident from the broader context of the passage just indicated:

> Although the relevant evaluations explicitly concern action, it seems that they reflect claims about the underlying reasoning. For instance, the relevant intuition in SUR-

[28] One special argumentative strategy worthy of note here is the kind of knowledge–action link posited by Jeremy Fantl & Matthew McGrath, according to which 'if you know that p, then p is warranted enough to justify you in Φ–ing, for any Φ. . . . [I]n saying that p is warranted enough to justify you, we are saying that no weakness in your epistemic position with respect to p—in your position along any of the truth-relevant dimensions—stand in the way of p justifying you.' Fantl and McGrath (2009, p. 66) As I mentioned in §5.2 above, their formulation of knowledge–action links is different from the reason-based version I favour; this is the reason why. I think Fantl & McGrath's principle is false for a reason that does not extend to KR. It seems to me that all of the following might be possible:

1. p is warranted enough to be a reason to Φ.
2. p is not a strong enough reason to justify Φ-ing.
3. the proposition that one knows that p would, if held, be a strong enough reason to justify Φ-ing.

Imagine a situation where you know that p, but fail to know that you know it, but where only the latter would be enough to justify action. (Perhaps ordinary high-stakes cases are like this; or if you prefer, you can set up artificial situations, like ones in which one has promised not to Φ unless one knows that p.) On the assumption that failure to know that one knows that p constitutes an epistemic shortcoming vis-à-vis p, this is a case where p is held as a reason, but where an epistemic shortcoming vis-à-vis p prevents you from being justified in acting. Brown's intuitions about the Surgeon case *are* inconsistent with Fantl & McGrath's knowledge–action principle, even though, as I have been arguing, they are not at all inconsistent with KR.

[29] Throughout the chapter, I use 'ethical' in a broad sense having to do with how one ought to act; this may not be exhausted by considerations of morality.

[30] Locke (2014, p. 83).

> GEON is that the surgeon should not rely on the premise that it is the left kidney which is affected in practical reasoning. For instance, she should not reason that since it is the left kidney which is affected, she should remove the left kidney straightaway rather than check first before operating. Brown (2008a, pp. 1145–6)[31]

The thing Brown calls 'the relevant intuition', which, as Locke points out, *is* in conflict with KR, is a theoretical step removed from the explicit evaluations. It is, it seems to me, invoked as an attempt to explain the more epistemically basic intuition, which concerns what action is appropriate. Although I disagree with Brown's particular attempt to explain this evaluation of action, I think that her invocation of this order of explanation is the right one. It's reasonable to be at least moderately confident in many of our intuitive judgments about whether one is, in a particular epistemic position, justified in undertaking a particular action; this is a very quotidian sort of evaluation to make. By contrast, I think we should be *not at all* confident in endorsing any intuitions we find in ourselves about whether one is justified, given a particular epistemic situation, in using a particular propositional content as a premise in practical reasoning. This highly theoretical notion is one about which any pre-theoretic intuitions we have are of minimal or no epistemic significance. (Speaking for myself, I just don't introspect any theory-independent intuition either way as to whether the surgeon is justified in using the proposition that *the left kidney is diseased* in her practical reasoning.)

What Brown calls 'the relevant intuition' is supported primarily by the considerations she mentions around it—namely, that it would be improper to operate without checking the chart. But as I have observed above, there are many possible explanations for that fact that do not require endorsing 'the relevant intuition'. There are other evaluative intuitions at play too, beyond the one involving the simple action. Brown gives one of them in the quoted passage above: 'she should not reason that since it is the left kidney which is affected, she should remove the left kidney straightaway rather than check first before operating'. I agree with this evaluation too, but it doesn't imply that she isn't permitted to use that proposition in her practical reasoning. One may dispute this reasoning without disputing that the fact in question is genuinely *a reason* that is held. That proposition is, for all Brown has said, just not a *good enough* reason to remove the kidney straightaway without checking. Remember again the case of *Fairy Law*: it is true that the Queen ought not to reason, 'since the defendant has been in a good mood, she should be banished'. But it would be a total *non sequitor* to infer from this that she's not permitted to use the proposition *that she's been in a good mood* in her practical reasoning. It *is* among her reasons; it just isn't a reason that would justify that particular action.

A similar dialectic plays out with respect to the putative counterexamples to the necessity of knowledge for reason. I pointed out above that in such cases—for example, in Gettier cases in which a subject seems appropriately to act as if he had knowledge—there is inevitably some knowledge still present, and the known facts seem still to support the reasonable action. For example, in Brown's potato case, I observed that the partner knew that a particular list indicated that they needed potatoes, and that this might well be a good enough reason to justify his buying potatoes, and to make his failure to buy potatoes criticizable. However, a more involved version of the argument might attempt to invoke subtler intuitions. One wouldn't merely cite the fact that he should have bought potatoes; instead, one might attempt to offer the more specific criticism that he should have used the proposition that *they are out of potatoes* as a premise in a bit of practical reasoning with the conclusion that he should have bought potatoes. Some of Brown's own remarks suggest this reading of her argument. Here is the broader context of her story about her parter's failure to buy potatoes:

> One way of reinforcing this conclusion is to consider a different form of criticism of action in which one criticises an agent for failing to act on a belief which is knowledge.

[31] This passage also appears on Brown (2008c, p. 177).

> For instance, in our earlier example, I criticise my partner for failing to get potatoes on his way home. In the example, he calls in at the supermarket to get beer but fails to buy any potatoes, despite knowing that we were out. I criticise his action saying: 'You should have got potatoes. You knew we were out'. It seems that this criticism would not be undermined if it turned out that he didn't know we were out of potatoes but was in a Gettier situation in which he merely had the justified true belief that we were out. [Gettier details omitted] Discovery of the kids' prank in no way mitigates the negative assessment. For instance, it would be ridiculous for my partner to try to defend his action by pointing out that he didn't know that we lacked potatoes, but only had a justified belief that we lacked them. Further, it seems that after the kids' tricks have been discovered, I can still criticise his action by saying, 'You should have got potatoes. You thought they were on the shopping list'. The defender of [KR] might attempt to deal with the example by claiming that we are confusing the assessment of the reasoning with that of the action or the agent. My partner's failure to buy potatoes is not the right action (we won't have anything to feed the kids), and he's surely blameworthy for getting beer but not potatoes. However, distinguishing these dimensions of assessment does not undermine the intuition that his reasoning is at fault: surely, given his evidence, he should have treated the proposition that we are out of potatoes as a premise in reasoning about what to do. He should have taken this proposition for granted in his practical reasoning. Brown (2008c, p. 173)

The virtue of this version of the argument is that it engages much more directly with KR. If Brown's intuition that *he should have used the proposition that they are out of potatoes as a premise in a bit of practical reasoning with the conclusion that he buy potatoes* is clearly correct, then his failure to know that proposition is in clear tension with KR. (On the assumption that one should only use propositions as premises in practical reasoning if they are reasons possessed, it is a literal counterexample.) But there are two dialectical drawbacks with relying on this kind of intuition. First, as highlighted above, this is a highly theoretical intuition, and not the sort of thing it's particularly mandatory to suppose is obvious. Second, it is false. Brown's suggestion here is that, even setting aside the apt criticism of the partner's action (coming home with no potatoes), he is also criticisable for a particular failure in reasoning. Brown thinks her partner did at least *two* things wrong:

> (11) He didn't buy potatoes.
>
> (12) He didn't use the proposition that *they are out of potatoes* as a premise in his practical reasoning.

To consider whether Brown is right to think that these critcisms *both* apply, we can consider versions of the case where he'd be guilty of one but not the other. For instance, suppose as before that he is in a Gettier case with respect to the proposition that they are out of potatoes, but that he does know that 'potatoes' appears on his list. Suppose that he goes to the supermarket, and uses the proposition that 'potatoes' appears on his list—*not* the proposition that they're out of potatoes—as a premise in a bit of practical reasoning, resulting in the action of buying potatoes along with his beer. In this case, it seems rather clear that his partner would have nothing to complain about. Even if she somehow discovered that his premise had been a claim about the list instead of one about the cupboard, she would have no cause for complaint.[32] This suggests that Brown is wrong to consider

[32] In §5.7 below I will dispute whether, in typical instances, there is a fact of the matter about which premise he's relying on, but we may suppose this is an extraordinary case in which he's explicitly and deliberately thinking of the proposition about the list.

(12) to be an appropriate criticism in a way separate from (11). So I don't think Brown is justified in claiming the very particular diagnosis she does of what goes wrong in cases like these; although it is clear that an action is warranted, which implies that one has *some* reason to perform it, we cannot in general assume it appropriate to criticise one for failure to rely on a *particular* content in practical reasoning. Even if one manages to have intuitions to the effect that one should engage in reasoning so specific, those intuitions ought not to be trusted.

Brown's version of the argument relied on intuitively acceptable criticisms of a subject in a Gettier situation. My remarks above apply equally well to versions of the argument like Mikkel Gerken's, that rely on positive ascriptions of rationality to subjects.[33] Gerken presents two similar cases—in one, a subject sees a lion on a hill; in the other, a subject is in a Gettier case with respect to the proposition that there is a lion on the hill. (He sees a lion-shaped rock on a hill that coincidentally occludes a hidden lion.)[34] Gerken says that it is 'overwhelmingly plausible that [both his characters] may use their respective beliefs that there is a lion on the hill in practical reasoning'. I don't find this at all obvious; what is obvious is only that in each case it is reasonable for the subject to take the precautions one would take if one knew there were a lion present. But this can be explained by invoking the subjects' knowledge, even in the Gettier case. For example, the known proposition that *there is something that looks like a lion on the hill* is a good reason for him to back away carefully.

5.7 The Thought-Bubble Model of Practical Reasoning

One point of pressure I have received against the suggestion of the previous section[35] is that one needn't rely on intuitions about what the subject is permitted to rely on; one could instead *stipulate* that the subject is relying on a particular content, even though it is unknown. If the intuition about such a case is that the subject is proceeding permissibly, the objection to KR is retained. Suppose we modify Gerken's case thus:

> As Ernest is walking across the plains of Masai Mara, he sees a rock that looks just like a big lion lying on a nearby hill. Consequently, he forms the belief that there is a lion on the hill. He uses this particular belief— the belief *that there is a lion on the hill* (rather than any weaker belief like the belief that it looks like there is a lion on the hill) as a premise in a bit of practical reasoning with the conclusion that he should load his rifle and slowly back away. As it happens a real lion lies hidden in the grass on the hill invisible to Ernest.

This version of the case *stipulates* that Ernest is relying on a proposition that is, we may agree, unknown. So if one finds it intuitively obvious that Ernest is doing everything right in this case, one has a counterexample to KR, on the assumption that one shouldn't rely on a proposition that isn't held as a reason. What should one think about stipulations like this one? I am suspicious of them. One way to motivate skepticism about them is to contemplate, what is the difference between the version of Ernest just proposed, and one who performs the same action for the reason that it *looks* like there's a lion on the hill? What is it about a subject in virtue of which he is relying on a Gettiered belief about a lion, rather than a knowledgable belief about an apparent lion? It is not obvious that the distinction tracks a genuine psychological difference.

A certain cartoon conception of practical reasoning might suggest there is a clear difference here. On the cartoon picture, decisions are inevitably the result of the manipulation of particular sentences

[33] Gerken (2011, pp. 535–6)

[34] I focus here on Gerken's negative project against the knowledge norm. In his broader project, Gerken means to motivate a competing approach to the knowledge first programme he finds preferable on broad theoretical grounds, including this one.

[35] For instance, from Mikkel Gerken in conversation.

in the head—the kind of things that would appear in thought bubbles. John Hawthorne and Ofra Magidor describe an 'admittedly crude model' of practical reasoning along these lines: 'People have an Aristotelian syllogism box in which various propositions appear as premises, instantiating practical syllogisms that yield action. A necessary condition for a proposition's being a motivating reason is that it figures as a premise of practical reasoning in this way'.[36] They go on to observe that in cases like Gerken's, the model in question would render it implausible that the subject is proceeding on the basis of a known fact.[37] We may call this the 'thought bubble' conception of practical reasoning.

On such a conception, there would be a very straightforward psychological difference; see Fig. 5.1.

Figure 5.1: According to the thought-bubble model, typical decisions fall determinately into one category or the other. Illustration by L. Syd M Johnson.

But as a description of ordinary psychology, this cartoon picture does not appear to be very realistic. When one performs a particular action, even a deliberate one like loading a rifle, one does act for reasons—but those reasons are not in general the things that one says to oneself in a speech-bubble-like way. I packed sunscreen in my bag today; this was a deliberate and thoughtful action; nevertheless, it would be a mistake to try to argue about which of the following were relied on as a reason:

s_1 It will be sunny this afternoon.

s_2 The forecast indicates that it will be sunny this afternoon.

s_3 It's been sunny most of the last several days.

s_4 It is summer.

o_1 I will be exposed to the sky this afternoon.

o_2 I will be outdoors this afternoon.

[36] (Hawthorne and Magidor, 2016, p. 6 (manuscript))

[37] Although they describe the picture in question as 'admittedly crude', they do not explain the respects in which they find it unlikely to be adequate. This is surprising, since the failure of the crude model to deliver the verdict that facts about appearances can be motivating reasons in such cases is their only stated argument against this idea, which otherwise fits well with the approach they defend. As a result, they are compelled to defend rational actions without any motivating reason. Hawthorne and Magidor (2016, p. 12).

o_3 I am going to a World Cup match this afternoon.

o_4 I have made plans to go to a World Cup match this afternoon.

All of these are among the things I believed;[38] none were things I thought-bubbled to myself in an explicit bit of reasoning that led to my taking sunscreen. There's just no psychological fact of the matter as to whether I relied on s_1 or s_2 as a reason. In at least the vast majority of ordinary cases of proper practical reasoning, subjects are responding rationally to their total body of reason; selecting a particular content as *the* reason for which they act is a post hoc theoretical project of interpretation.[39] The thought-bubble model of practical reasoning does not capture ordinary reasoning.

I don't go so far as to claim that the thought-bubble model is *never* an accurate description of reasoning. I can imagine, for example, that one might establish a kind of artificial intelligence that leaves a record of the particular propositions upon which it is relying. It may even be that some humans occasionally reason according to the thought-bubble account—*perhaps* this happens in cases when we are being very deliberate in System 2 decision-making (though even then, I am skeptical). But this is far from the norm; we should not expect to have clear intuitions about a version of Ernest who is determinately relying on the content *there is a lion* rather than *it looks like there's a lion*. So we shouldn't draw conclusions from such heavily and theoretically stipulated cases.

This kind of strategy is also my response to Hawthorne's challenge to contextualist implementations of KR, mentioned in §5.3 above. Recall the alleged implication of such a view that the following may be true in contexts more skeptical than Strephon's:

> (5) Strephon doesn't know that Celia is his Aunt, but he is perfectly reasonable in embracing her for the reason that she's his Aunt.

The contextualist can easily apply the apparatus discussed in the preceding sections to this challenge. Relative to the skeptical contexts in which the first conjunct in (5) is true, 'that Celia is his Aunt is the reason for which Strephon embraces her' is false, and so the second conjunct in (5) is also false. What *is* true, and what allows us to capture the intuition Hawthorne thought would motivate (5), is that although Strephon doesn't 'know' that Celia is his aunt, he is acting perfectly reasonably in embracing her—but the reason we cite for his embracing isn't *that she's his aunt*; it's something weaker that Strephon *does* know (perhaps that Iolanthe *told* him that she's his aunt). The thought-bubble model of practical reasoning might charge the move as a kind of ad hoc reinterpretation. But if, as I've suggested, the thought-bubble model is false, this is no reason for dissatisfaction.[40]

5.8 Counter-Closure

The dialectic of the previous section is very closely related to that in the epistemological literature on 'counter-closure'—the plausible idea that gaining knowledge by inference requires knowledge

[38] Other of my beliefs, such as the belief that sunscreen protects one from the sun, are also relevant.

[39] So in at least this limited way having to do with deliberate practical reasoning, I agree with the content indeterminacy of Dennett (1987, p. 40). My approach fits naturally with the suggestion of Fogal (2016) that metaethical emphasis on particular *reasons* exaggerates their importance. Fogal's idea, with which I have considerable sympathy, is that one's *total body of reason*, as typically expressed by the mass noun 'reason', is more fundamental than are entities denoted by the count noun 'reasons'.

[40] The thought-bubble model is also sometimes assumed in theoretical reasoning; (McGlynn, 2014, p. 73) cites Joyce (2004, pp. 301–2), arguing, against E=K, that when confronted with misleading testimony, subjects use false testimonial contents as evidence. A defender of E=K will likely hold that there is other, weaker evidence (such as the fact that the testimony occurred) available; Joyce appears to reject this view on the grounds that the subject is determinately relying on the false content, not the known fact of testimony. Only on a thought-bubble model of theoretical reasoning does this stipulation seem to me to be plausible.

of the premises.[41] The typical objections to counter-closure tend to involve cases like this one from Ted Warfield:

> *Handout*
>
> Counting with some care the number of people present at my talk, I reason: 'There are 53 people at my talk; therefore my 100 handout copies are sufficient.'. My premise is false. There are 52 people in attendance—I double counted one person who changed seats during the count. And yet I know my conclusion. Warfield (2005, pp. 407–8)[42]

As Brian Ball and Michael Blome-Tillmann have pointed out, however, on psychologically plausible versions of this case, the false specific belief needn't and shouldn't be thought of as *the* premise in the subject's reasoning. As they write,

> it should already seem mildly suspicious that, in the explicit description of the case, *only* the belief that there are 53 people in the room is mentioned as supporting the (justified) belief (indeed, knowledge) that Warfield's 100 handouts are sufficient; for surely the fact that 100 is more than 53 is *also* relevant. That is, in order to infer that his 100 handouts are sufficient, from the premise that there are 53 people in the audience, Warfield must be in some way sensitive to the fact that 53 is less than 100. We mention this point not to suggest that in *Handout* Warfield does not know his conclusion (by virtue of failing to base his belief that he has sufficient handouts on this further premise), but rather to suggest that in the case of *Handout* there are almost certainly some *further* albeit *tacit*, or *subconscious* beliefs present, which enable him to undertake the inference. And once this is recognised we think that it is fairly clear what is going on in the case at hand: the false belief that there are 53 people in the room is *epiphenomenal* as far as the causation and the justification of the belief that the 100 handouts are sufficient are concerned; instead, the latter belief is causally grounded in, and informationally sensitive to, certain other, tacit or subconscious beliefs.
>
> More specifically, we think that our intuition that the subject of *Handout* knows his conclusion is tracking certain features of the case which are not spelled out explicitly, but which are nonetheless projected by readers as present. In particular, we think that what is really going on in *Handout* is something like the following: Warfield knows (perhaps tacitly) that the result of his count is '53'. He then forms two further beliefs: one is the explicit belief that there are 53 people in the room; the other is the tacit belief that there are approximately 53 people in the room. He then employs the latter belief, together with his knowledge that 53 is less than 100, to conclude that his 100 handouts are sufficient. The last belief ... is based upon other beliefs which also constitute knowledge, and so is no counterexample to counter closure. Ball and Blome-Tillmann (2014, pp. 554–5)

Ball and Blome-Tillmann's observation about apparent counter-closure counterexamples seems to me broadly correct; only if one assumes a psychologically implausible 'thought-bubble' approach to deductive inference is one entitled to stipulate that subjects are relying on false beliefs, as opposed to other true ones in the neighbourhood, in their reasoning.[43] And if one were to make stipulations strong enough to require the thought-bubble interpretation—if one stipulated, for instance,

[41] See also the similar dialectic with respect to the factivity of evidence, challenged by Rizzieri (2009, p. 237).

[42] For similar cases, see Luzzi (2010), Luzzi (2011), and Fitelson (2010).

[43] I do have one incidental point of disagreement with Ball and Blome-Tillmann; although I agree with all their central points here, I am not convinced that the proposition that *53 is less than 100* need be among the subject's reasons; I think that

that Warfield rather bizarrely remains carefully agnostic about the proposition that his counting procedure delivered the result '53', relying instead *solely* on the proposition that there *are* (exactly) 53 people in the room—then it is far from clear that his process yields knowledge.

Human reasoning, in both the practical and the theoretical realms, is psychologically complex; the thought-bubble model does not begin to do it justice. While it is useful for many purposes to ascribe action on the basis of particular reasons, we should not forget that this is an interpretive, rationalising project. If KR is right, then it is a constraint of that project that one only attribute pieces of knowledge as reasons; if the broad thrust of the last several sections is right, this constraint isn't a terribly onerous one, or one with radically counterintuitive consequences.

All of this depends, however, on my assumption that in the kinds of cases we've been considering, the known propositions stand in suitable rationalising relations to the actions in question. As I have been emphasizing, whether this is so depends on substantive ethical questions about what reasons would justify what actions. It is now time to consider some particular attempts to argue that plausible ethical theories are inconsistent with my suggestions.

5.9 Locke on Ethical Theory

In his (2014) response to Ichikawa (2012), Dustin Locke argues that a kind of Bayesian decision-theoretic framework, combined with certain assumptions about Brown's surgeon case, demonstrates the latter to be a counterexample to KR after all.[44] Locke grants for the purpose of argument (in this part of his paper) my contention that Brown's argument is only convincing if we assume that the proposition that the left kidney is diseased would be a sufficient reason to justify beginning surgery without checking the chart; he goes on to argue that, contrary to my suggestion, this assumption *does* follow from the correct ethical theory.

Locke constructs a version of *Surgeon* where, if one assumes that (a) a Bayesian picture of practical rationality is correct and (b) practical reasons talk translates into the Bayesian talk by letting one conditionalise on one's reasons, we can derive Brown's assumption mentioned above. Although I consider both of Locke's assumptions very debatable, I don't propose to argue on that score at present. For there is a more interesting response to Locke's argument available as well: the case he tries to stipulate is on reflection much more problematic than he assumes. Locke makes several detailed theoretical stipulations. He writes:

> To see that there is such a way of filling in the details, suppose that all of the following are true.
>
> 1. The surgeon cares about, and only about, whether the patient lives.
> 2. The surgeon has credence 1 that exactly one of the patient's kidneys is diseased, and a .99 degree of credence that it is the left kidney.
> 3. If the surgeon performs the surgery without first checking the chart, she will begin it immediately; if she first checks the patient's chart, she will begin the surgery in one minute.
> 4. The surgeon has credence 1 that were she to check the chart, she would then remove the correct kidney.

one can reason directly from there being approximately 53 people to there being fewer than 100 people, without the need for a mathematical premise. On my view, such a premise is redundant in the way gestured at by Carroll (1895). Ichikawa and Jarvis (2013) offers some reasons for treating arithmetical claims in this way.

[44]The idea that Bayesian approaches stand in tension with reason-based approaches like KR has been developed by e.g. Douven (2008) and Schiffer (2007) (cited by Weisberg (2013), who gives a response different from mine).

> 5. If the patient has the correct kidney removed during the operation, then there are the following probabilities that he will live, depending on how soon the surgery begins: (5a) If the surgery begins immediately and the correct kidney is removed, there is a probability of 1 that the patient will live; (5b) If the surgery begins in one minute and the correct kidney is removed, there is a probability of .999 that the patient will live.
> 6. If the patient has the wrong kidney removed during the operation, then the probability that he will live is 0.
>
> Since, as Brown specified, the surgeon has spent the morning with the patient discussing the upcoming surgery, it is quite plausible that the surgeon knows that the left kidney is diseased, despite the fact that she does not have credence 1 that the left kidney is diseased. Locke (2014, pp. 86–7)

Locke goes on to observe that given these stipulations, the expected value of checking the chart first is higher than that of operating directly—and so concludes that she should check the chart first. But he also observes that if one conditionalises on the claim that the left kidney is the correct one, operating directly has a higher expected value, and claims on this basis that KR makes the wrong prediction: if KR were right, then the surgeon should conditionalise on her knowledge that the left kidney is the one to remove, and so she shouldn't check first.

Locke's stipulations, however, are problematic for at least two reasons. First, the assumption that a .99 credence is consistent with knowledge is dubious. It is not at all innocent to assume that knowledge is consistent with non-maximal credence like this. For lottery-related reasons, Locke is probably committing himself to the denial of multi-premise closure here.[45] Similarly, there are good reasons to deny that the surgeon believes (outright) that the left kidney is diseased under these circumstances, either for conceptual/metaphysical reasons[46] or for psychological reasons.[47] If this is right, then Locke's argument commits to knowledge without belief.[48]

My second worry concerns stipulation number 1: that this is a surgeon who cares *only* about the life of the patient. From a realistic point of view, this is a very strange surgeon. According to Locke's stipulations, the surgeon cares nothing at all about any of the following: whether she follows hospital procedure; whether she sets a good example for the students observing; whether she is proceeding rationally; whether she makes herself and her employer vulnerable to a possible malpractice suit; whether she acts only on propositions that she knows. These strong assumptions are not idle; if we allow that she cares about any of these things, the utility calculus will not require her to go without checking, even when she conditionalises on the content of her knowledge that the left kidney is diseased. (Suppose she cares about whether she acts only on that which she knows, and that she doesn't know whether she knows; then in operating without checking, there is a substantial risk of the negative outcome of acting on something she doesn't know.) But these very strange assumptions will make our intuitions harder to trust. When we try to imagine ourselves in her position, we naturally assume she cares about the ordinary things people might care about. Stipulating that she only cares about one thing—not even mentioning the many other things we have to remember to disregard—makes it very hard to get into her mindset. So I'm inclined to mistrust intuitions about so heavily-stipulated a case.

[45] Indeed, for reasons like the ones Maria Lasonen-Aarnio (2008) has emphasised, he may very well commit himself to denying *single*-premise closure.

[46] See e.g. Weatherson (2005); Clarke (2013)

[47] See e.g. Nagel (2010)

[48] For more on outright belief, see Ch. 7.

5.10 Schroeder on Ethical Theory

According to KR, only what is known can be among a subject's reasons. I have suggested that in many would-be counterexamples to the necessity of knowledge for reason, where propositions thought to be the central reasons aren't known, *other* known propositions can provide reason enough to justify the actions in question. But it's not enough for me to point to *some* propositions that are known, in order to show that KR delivers the correct verdict. Those propositions must be propositions that would justify the actions in question. The dialectical situation is somewhat subtle here; insofar as I am attempting to resist particular *arguments* against KR, it is enough to show that they assume without argument non-obvious ethical verdicts. If I have succeeded in showing that the arguments require assumptions that haven't been defended, then I succeed in showing that their arguments are inconclusive. Still, it would be desirable for the defender of KR to do something stronger—to demonstrate that the assumptions in question are actually false. The obvious way to do this would be to defend a particular theory about what reasons rationalise what actions. I'll sketch the framework for such an approach in §5.11; for now, however, I'll focus on the negative project of showing that the counterexamples to KR are inconclusive. It is helpful in this context to remember that KR is an approach with antecedent appeal; everyone needs some epistemic story about which contents get to be among a subject's reasons for action; KR is a natural and attractive candidate. So if one can show that the arguments against KR are inconclusive, one may restore the presumption KR enjoys.

However, I have encountered one kind of schematic argument against the permissibility of some of the kinds of moves I've suggested. Let's take another example of a case where one seems to act reasonably based on a false belief.

> *Buttercups*
>
> Celia plays a trick on Fleta, falsely telling her that the Queen has commanded that she spend the afternoon collecting buttercups. Fleta spends the afternoon collecting buttercups.

If we suppose that Celia's trick is well-played, and that the idea that the Queen has issued such a command is plausible enough, Fleta would be reasonable in spending her afternoon among the buttercups, even though her false belief that the Queen commands it is playing a significant role in the explanation of her action. I suggested in §5.4, however, that the defender of KR has no difficulty accommodating this verdict. Although one cannot cite *that the Queen has commanded it* as a reason that explains why her action is reasonable—since it is false and therefore unknown—there are other known propositions that can be cited. For instance, *that Celia told Fleta that the Queen has commanded it* is among Fleta's reasons. The question before us now, however, is whether this proposition is a reason that would justify Fleta's action of collecting buttercups.

An argument similar to one offered by Mark Schroeder suggests that it wouldn't.[49] Schroeder considers a move somewhat similar to the one I have been advocating with respect to the case of Bernie and his martini glass full of gasoline, discussed in §5.2. That case is like *Buttercups* in that the agent has a false belief that one might think nevertheless makes a particular action (in this case, taking a sip) justified. Schroeder considers the idea that the reason that rationalises this action is *that Bernie believes his glass contains gin*.[50] Schroeder argues that this proposition isn't a reason to take

[49] I have also received something like this argument in conversation with Maria Alvarez and Clayton Littlejohn.

[50] As my own choice of examples may suggest, I do not prefer to cite psychological facts about what a subject believes as reasons for actions in cases like these; I do not think in general that the fact that one believes that p is a very good reason to do the things that p would be a good reason for, except insofar as, assuming a person is reasonably reliable, that one believes p is some evidence for p. Even in such cases, this kind of consideration does not typically seem to articulate an *original* source

a sip; so observing that he has it doesn't help to make taking a sip rational.[51] The reason it doesn't, Schroeder suggests, is that it fails what we might call the 'fully informed and beneficent bystander' test:

> Suppose that a fully informed and beneficent bystander is tallying pros and cons of Bernie's taking a sip. He's just noted the fact that the glass is full of gasoline in the cons column. Does he now reflect, "but on the other hand, at least there's this much to say for it—Bernie *believes* that the glass contains gin and tonic"? This seems like a strange thing to say. Bernie's taking a sip is no better of an idea, just because he is in the dark. Now suppose that Bernie were to find out that he believes that the glass contains gin and tonic. Would that be the sort of thing to settle him on what to do—on taking a sip? On the contrary, Bernie would just as soon not drink gasoline, even if he believes that it is gin and tonic. When he is deliberating about what to do, what he believes matters to him only if it is an indication of how things actually are. Neither of these characteristics of reasons for someone to act seems to be present in Bernie's case. This suggests that the fact that he believes the glass to contain gin and tonic is not itself a reason for him to take a sip. Schroeder (2008, p. 62)

Schroeder's seems to be offering this argument:

1. A fully informed beneficent bystander wouldn't put *that Bernie believes his glass contains gin* in the 'pro' column for taking a sip.
2. Reasons to Φ are all and only those things a fully informed beneficent bystander would put in the 'pro' column for Φ. Therefore,
3. *that Bernie believes his glass contains gin* isn't a reason to take a sip. So this reason, even if had, wouldn't rationalise the action.

Something like this argument might also be marshalled to argue that in my *Buttercups* case, *that Celia told Fleta that the Queen had commanded her to collect buttercups* is no reason under the circumstances for Celia to collect buttercups; a fully informed beneficent bystander would know full well that this testimony was misleading, and so would not enter it into the 'pro' column for collecting buttercups. But I do not accept Schroeder's assumption that only things a fully informed beneficent bystander would put in the 'pro' column for Φ-ing are reasons to Φ. Schroeder offers no defence of this assumption; indeed, he doesn't even articulate it. His assumption seems to be that it is obvious. But it is strongly at odds with a very natural picture of how we ought to reason under uncertainty. Indeed, there is something slightly perverse about the project of drawing up a pro and con list for a *fully informed* beneficent bystander. Such lists are paradigmatically useful for subjects

---

of justification for action. Maybe there are some deviant cases where one has no idea why one forms a belief, but trusts oneself about it anyway and so acts on the fact of that belief, and does so reasonably, but this is not the usual case. (Suppose an amnesiac finds himself with the belief that a certain box is dangerous; this might be reason for him (or indeed anyone else) not to open it.) For related discussion, see Ichikawa and Jarvis (2013, pp. 271–2). This is why I tend to focus, not on the psychological fact of the belief itself, but on the (typically external) facts that rationalise the belief. It is surprising that Schroeder does not consider any such candidate reasons—*that he asked for a gin and was given this drink*, for instance—the psychologistic one is the only known proposition he considers as a candidate reason. Still, the objection Schroeder gives to the belief move looks at least potentially to be applicable to my preferred version of the move as well.

[51] In fact, Schroeder also argues that this proposition might not be among the reasons he even *has*. See Schroeder (2008, p. 62). While I'm unconvinced on this score, I don't think it's where the action is, in part because I don't prefer to cite psychologistic facts in such cases as reasons anyway. See previous footnote.

making decisions in less than fully informed positions.[52] If you knew all the relevant facts, you wouldn't need a pro and con list.

5.11 Reason-to

So I don't think that Locke's or Schroeder's particular approaches to the *reason-to* relation succeed in causing unavoidable trouble for the moves I've been making on behalf of KR. But this is not to say anything in particular about what it *does* take for a given reason to be a reason to perform a particular action. In this section I'll briefly sketch a few possible approaches to this question, and suggest that they seem friendly to the approach I've been developing. In the following sections, I'll consider how, when combined with KR, they bear on broader questions about practical reasoning and contextualism.

It is helpful in getting started in this context to consider the parallel dialectic about evidence. Evidence for q can be both inconclusive (it doesn't entail q) and misleading (q turns out to be false); that doesn't prevent it from being genuine evidence in favour of q. One theoretical approach to evidence that accommodates this theory is that defended by Timothy Williamson: there is an antecedent evidential probability function P, and evidence e is evidence for proposition q just in case $P(q|e) > P(q)$. That is to say, evidence is evidence *for* q just in case the probability of q given the evidence is higher than q's antecedent probability; colloquially, if it *raises the probability* of q.[53]

One way to characterise what it is for a given reason to be a reason *to* Φ characterises it as a particular kind of evidence. Stephen Kearnes and Daniel Star have defended a view in this vein:

> Necessarily, a fact F is a reason for an agent A to Φ if and only if F is evidence that A ought to Φ.[54]

---

[52] Benjamin Franklin is credited with first describing such pro–con lists, in a 1772 letter to Joseph Priestly, who had asked him for advice on a difficult question. As Franklin points out, making such lists is useful precisely when one is relevantly uncertain:

> 'When these difficult Cases occur, they are difficult chiefly because while we have them under Consideration all the Reasons pro and con are not present to the Mind at the same time; but sometimes one Set present themselves, and at other times another, the first being out of Sight. Hence the various Purposes or Inclinations that alternately prevail, and the Uncertainty that perplexes us. To get over this, my Way is, to divide half a Sheet of Paper by a Line into two Columns, writing over the one Pro, and over the other Con. Then during three or four Days Consideration I put down under the different Heads short hints of the different Motives that at different Times occur to me for or against the Measure. When I have thus got them all together in one View, I endeavour to estimate their respective Weights; and where I find two, one on each side, that seem equal, I strike them both out: If I find a Reason pro equal to some two Reasons con, I strike out the three. If I judge some two Reasons con equal to some three Reasons pro, I strike out the five; and thus proceeding I find at length where the Balance lies; and if after a Day or two of farther Consideration nothing new that is of Importance occurs on either side, I come to a Determination accordingly.' Franklin (1956, p. 26).

If, unlike Bernie, one knows that refusing to drink would have the best results, one would simply enter that fact into the 'cons' register; this would swamp anything that manages to get into the 'pro' column—perhaps, for instance, the fact that it would be polite to drink what is offered.

[53] Williamson (2000, Ch. 10). This is not to deny that many important questions about this framework remain; for example, how much background knowledge is encoded into the probability function P, and how we can make sense of there being evidence in favour of that which is already evidence. Still, the general framework is clear enough to help shed light on practical reason.

[54] (Kearns and Star, 2009), (Star, 2015, p. 13). This needn't imply that subjects are always able to recognise whether something is a reason for them to Φ—one may be unable to tell that something is evidence that one ought to Φ. (That someone will suffer unless you Φ is, I think, a moral reason to Φ, whether or not one has the appropriate moral sensitivity to this fact.) So I am inclined against a close association between being a reason to Φ and being the sort of thing that is likely to convince someone to Φ—see e.g. Manne (2014, p. 97). (But whether the kinds of connections proposed by 'internalists' like Manne are ultimately tenable depends on many details, such as the nature of the idealisation involved.)

This 'reasons as evidence' view predicts rather well the suggestions I made about Brown's and Gerken's cases above. Recall, I am trying to vindicate intuitions such as these:

- In Gerken's 'lion' case, the fact that there is something over there that looks like a lion is a reason to back slowly away.
- In Brown's 'potato' case, the fact that the list indicated 'potatoes' is a reason to buy potatoes.
- In Brown's 'surgeon' case, the fact that the left kidney is diseased is not a *conclusive* reason to remove the left kidney without checking the chart.

With respect to the lion intuition, the 'reasons as evidence' view requires that the probability that one ought to back away, given that that something over there that looks like a lion, is higher than the antecedent probability that one ought to back away. This is very plausible. If you start out in ignorance about whether there is anything that looks like a lion there, your credence that the subject ought to back away certainly ought to go up upon learning that there is something that looks like a lion there. Just the same goes for the potato case, which has the same structure.

The surgeon case is less straightforward to adjudicate; we need to make a choice about how to encode, within the reasons as evidence framework, the notion of a *conclusive* reason. (It is uncontroversial that the fact that the left kidney is diseased raises the probability that one should remove the left kidney without checking the chart; it is likewise uncontroversial whether it is *a* reason to do so. The key question is whether it is a *sufficient* reason to do so.) One natural implementation of the idea would have it that r is a *conclusive* reason to Φ just in case the probability that one ought to Φ, given r and the information encoded into one's background probability function, is 1.[55] But the discussion in §5.9 above raises reason to doubt that the probability that the surgeon ought to remove the left kidney without checking, conditional on the left kidney's being diseased, is 1. (For example, it is plausible to suppose that she would be subject to censure, perhaps even risking losing her job, if she did so.) There may well be different ways to think of conclusive reasons on the 'reasons as evidence' model, but I do not see any obvious reason to think that they would be inconsistent with my suggestions here.

A different view in a similar spirit would identify reasons, not with a particular kind of evidence, but on a more direct *analogy* with evidence. Recall: on the Williamsonian approach to evidential probability, evidence-for is probability-raising. In just the same way, we can suppose in the practical realm that reasons to Φ are reasons that raise the expected value of Φ-ing. We can posit an objective expected value operator analogous to Williamson's probability operator, and hold that reason r is a reason to Φ just in case $\mathrm{EV}(\Phi|r) > \mathrm{EV}(\Phi)$. The antecedent expected value of taking a coinflip bet with even payoffs is 0, since I'm equally likely to win as I am to lose. But if it will also amuse my friend, that fact will raise the expected value into positive territory, and so is a reason to take the bet. (The same would go if the initial EV were at some value other than 0, in which case that extra reason might or might not make a difference as to what I ought to do.)

This view mimics the reasons as evidence view with respect to the cases mentioned above. That there is something that looks like a lion very plausibly does raise the expected value of backing away; that the list says 'potatoes' raises the expected value of coming home with potatoes. And the fact about the kidney raises the expected value of operating without checking—but that's fully consistent with the preponderance of reasons favouring checking first.

The literature on reasons is vast, and I have engaged it only superficially here. It is not my project to stake out a substantive ethical theory about what reasons support what actions. But this sketch of such a view, and how it handles the cases I've been discussing, is I hope at least somewhat useful

[55] (Kearns and Star, 2009, pp. 235–6) commits to the claim that conclusive reason to Φ entails that one ought to Φ.

in bolstering my case that the verdicts I've been relying on about reasons in particular cases may ultimately be defensible.

5.12 Internalism and External Redundancy

So far, I have been emphasising that arguing against KR by citing intuitions about cases is more challenging than it's is often realised to be. Since KR only has direct implications about what reasons a subject has, and doesn't itself imply anything about which reasons would justify which actions, pairs of intuitions about knowledge/ignorance and appropriate/inappropriate action will never themselves comprise a challenge to KR; one would need to answer substantive ethical questions about the relationship between reasons and appropriate actions. I have also argued that promising approaches to these substantive ethical questions will not vindicate the extant challenges to KR—they do not show that KR has implausibly externalist implications. In this section, I will consider the broader relationship between KR and *internalism* about justified action. A plausible and attractive implementation of KR is not only *consistent* with internalism; it might even predict and explain it.

As in the case of Ch. 4, I am using 'internalism' in its supervenience sense. Internalism about warranted action amounts to the claim that facts about which actions are warranted are settled by a subgroup of factors 'internal' to the subject. As in the previous two chapters, there are various ways that one might demarcate the category of the 'internal', depending on one's broader epistemic commitments. In the language of Ch. 3, the central question concerns with what features we want to identify a subject's 'cognitive home'.[56] On one traditional version of internalism, the thesis would be that *intrinsic duplicates* will be equally justified in their corresponding actions. On a view that fits better with the neo-Moorean disjunctivist approach to perception discussed in Ch. 3, external-world propositions that are directly perceived will also be included among the 'internal'; the relevant statement of internalism about warranted action would be that the subject's basic knowledge (including basic perceptual knowledge of the external world) and cognitive processing together determine which actions are warranted. I describe this as a kind of internalism because it says that external factors to which the subject has no direct access can make no difference for whether an action is warranted. (So whether the shopping list is genuine or an excellent forgery does not bear on whether one ought to buy potatoes.) The notion of the 'external' is susceptible to the same diverse treatments—for any given understanding of the 'internal', the 'external' is that which isn't 'internal'.

So internalism about warranted action is, to an approximation, the claim that no purely external difference can make for a difference in what a subject ought to do.[57] Internalism follows from KR and this principle:

> *ER* Any time S's total knowledge K justifies action Φ, S's internal state is such that anyone in that state will have a body of knowledge K_i , which *itself* suffices to justify an action corresponding to Φ.

This is a claim to the effect that the *e*xternal is *r*edundant. Since knowledge is factive (and for Gettier-related reasons), one's total knowledge depends on features of one's external environment. In a good case, one might know that one needs potatoes, whereas in the bad case, one fails to know

[56] Again, I do *not* mean, as Williamson does, to identify cognitive homes with luminous aspects of experience; rather, cognitive homes are the domains over which we enjoy special responsibility. See §4.9.

[57] The reason this is only an approximation is content externalism; I assume that actions are individuated using contents which themselves may be externally (including socially) determined. So for instance, I may be justified in feeding my dog Mezzo; my intrinsic duplicate is not justified in feeding Mezzo—instead, he is justified in feeding another dog, i.e., his dog (who presumably looks quite like Mezzo).

this, even though the cases are internally (in whichever sense of 'internal' we prefer) alike. But although what total knowledge one has depends in part on one's external environment, ER has it that the additional knowledge one has in the good case is *redundant* in the sense that it doesn't make a difference as to what one ought to do—the knowledge that one continues to have, even in the bad case, is enough to warrant the preferred action shared in both cases.[58] I am not certain whether ER is true; consequently, I am not certain whether someone who accepts KR should endorse internalism about warranted action. I do think I've shown in this chapter that the *particular cases* that have convinced some theorists that KR implies externalism show no such thing. On reflection, I've never encountered a case that convinces me that KR fits poorly with externalism. When thinking through cases, it strikes me as very plausible that some kind of internalism about warranted action is correct, and that the kind of dialectical characteristic of ER works well. Although carefully adjudicating the question of internalism would require a much more detailed investigation of just how particular reasons warrant particular actions—if the suggestions of the previous section were on the right track, this would involve substantive study of the evidential probability and/or expected value operators—insofar as internalism about warranted action is plausible, that provides some reason to suppose that ER too is correct.

In the next section, I'll consider a relevantly parallel dialectic in the realm of consequentialist ethics; in §5.14, I'll turn to an objection to internalist versions of KR.

5.13 An Ethical Analogy

Consider an objection that James Lenman has pressed against consequentialist theories of morality, according to which one ought to perform that act which will have the best consequences. Lenman observes that in many cases—perhaps in almost all cases—agents are irredeemably ignorant of what the ultimate consequences of their actions (or potential actions) will (or would) be. His central cases are ones that impact which people exist in the world. These include decisions about whether to kill or whether to procreate, but also those decisions that might causally impact such decisions—i.e., nearly all decisions. (For example, whether I smile at a stranger this afternoon might have causal implications on her mood this evening, which may in turn bear on her future reproductive profile.) Since the effects of such issues magnify over the generations, and since the many future possible individuals whose existence hangs in the balance might well, if they do come to exist, make decisions of tremendous positive or negative moral consequence, the project of making even an educated guess about the total consequences of my action here and now is, according to Lenman, a foolhardy one. As he puts it:

> We do not have a clue about the overall consequences of many of our actions. Or rather—for let us be precise—a clue is precisely what we do have, but it is a clue of bewildering insignificance bordering on uselessness—like a detective's discovery of a fragment of evidence pointing inconclusively to the murderer's having been under seven feet tall. We may not be *strictly* without a clue, but we are *virtually* without a clue. The trouble for consequentialism then is that the foreseeable consequences of an action are so often a drop in the ocean of its actual consequences.' Lenman (2000, p. 350)

[58] Aidan McGlynn seems to be overlooking the possibility that the difference in knowledge may be redundant when he writes of a parallel case (involving theoretical rather than practical reasoning) that

> [t]he factors that determine whether one's attitudes are rationally held supervene on one's non-factive mental states. Call this thesis *rationality internalism*. Moreover, it's generally regarded as a platitude that rational subjects respect their evidence. It seems to follow that one's evidence must supervene on one's non-factive mental states too. McGlynn (2014, p. 67)

But this doesn't follow if the difference in evidence is redundant.

Insofar as it is an implausible implication of an ethical theory that it implies that we have virtually no clue what the morally right action is, Lenman considers this a serious problem for consequentialism.[59] My sympathies, by contrast, are largely with those of Alistair Norcross. In his (1990), Norcross imagines a dictator who invites you to decide whether or not Smith will be killed. Your decision is authoritative with respect to Smith—he lives or dies at your decision—but there are also other factors outside of your control or ability to predict. In particular, a coinflip ties your decision to the fate of some other number of prisoners.[60] If the coin lands heads, they will share Smith's fate (i.e., they live if and only if you decided that Smith lives); if it lands tails, they'll have the converse fate (they die if and only if you decided that Smith lives). So if the coin lands heads, it's clearly better to spare Smith, since that means sparing everyone; if it lands tails, it's better to kill Smith, since that's the way to spare everyone else. Since one has no ability to predict how the coin will land, a version of Lenman's argument might charge that one has no hope of approaching the dilemma on consequentialist grounds. But as Norcross points out, 'if you have no more reason to think that a choice of death for Smith will benefit [the others] than a choice of life for Smith, and you know which will benefit Smith, you choose life for Smith. In this case, the unforeseeable consequences of your choice are greater than the foreseeable consequences, but the basis of your choice is still clear.'[61] I find Norcross's observation here very plausible. Although there are often many unforeseeable consequences of great importance, there is no reason one can't nevertheless make ethical decisions in a rational way on the basis of the *foreseeable* consequences.

The central move here is a version of the same one I've been emphasising in defending an internalist version of KR. Subjects in cases who lack what look like important bits of knowledge nevertheless have *other* knowledge that is of significance. Ernest doesn't know there's a lion out there, but he knows that there's something that looks like a lion out there. And yes, whether there's a lion out there is a much more important question than whether there's something that looks like a lion out there—in just the same way as the question of whether a particular decision would spare seven million lives is much more important than whether it will spare Smith's life. Possession of the more important knowledge would render the less important knowledge practically irrelevant, in either case. Nevertheless, in both cases, even lacking the more important knowledge, subjects can act on that which they do know.

5.14 A Challenge to Internalist KR

I hope that so far I have achieved two things: first, to argue that there's no obvious reason that KR is inconsistent with internalism, and second, to at least give some motivation for the idea that the combination of these two theses is *prima facie* attractive. In this section, I consider a challenge for this combination of views.[62]

Here, schematically, is the issue. Suppose that the totality of a subject's reasons count against a particular action, but *only barely*—if there were *slightly more* reason to perform the action, or *slightly less* reason not to, then the action would be justified. Now modify the case in a purely external way that undermines some of the knowledge that counts against the action, or that creates new knowledge in favour of it, but that leaves knowledge of the opposing reasons unaffected. If KR

[59] This is most directly a problem for so-called 'objective consequentialism', which has it that the morally best action is that which *would actually* have the best consequences. Burch-Brown (2014, p. 108) argues that a version of the worry extends to 'subjective consequentialism', according to which one should choose the action with the highest expected value on one's evidence. For the distinction, see Jackson (1991) and Howard-Snyder (1997).

[60] Norcross gives three versions of the case, with one, two, or seven million other people at stake.

[61] (Norcross, 1990, pp. 254–5)

[62] Several people, including Timothy Williamson, Daniel Greco, and Maria Lasonen-Aarnio, have each pressed a version of this challenge on me in conversation.

is correct, such a change must, the objection goes, to make the difference in warranting the action. For it would alter the reasons in an asymmetric way. It would remain true that *there are* reasons against the action, but those reasons are bound to be weaker than the reasons that are now present in favour of it. And since they were initially *very nearly* able to warrant the action, they're now (barely, perhaps) able to do so.

It is difficult to state an instance of this schema with confidence, because it relies on a difference between cases in which the reasons are only very slightly in favour of some action and ones in which they're slightly opposed; typical cases will be difficult to recognise as such. But kinds of cases involving pure symmetry might be more easily recognised. Consider for instance this pair of cases:

> **Fairy Path (Good)** Phyllis wants to find a fairy, but doesn't care which one. She knows that some fairies came down this way, but has no idea whether each one went left or right at this fork in the road. (The fairies make their decisions independently.) Until she learns more about the fairies, her total reasons support taking each path to an equal degree. But then she gets some new information. Mountararat tells her that Celia the fairy recently said she usually takes the left path. Phyllis now has some new reason to take the left path. Then Tolloller tells her that Leila the fairy recently said *she* usually takes the *right* path. This is reason for Phyllis to take that path. Tolloller and Mountararat are each expressing knowledge they have, and Phyllis comes to know both contents. Since Tolloller and Mountararat are equally reliable and trustworthy, and since Phyllis is indifferent as to which fairy she finds, her total reasons should still be symmetric with respect to the two paths.
>
> **Fairy Path (Bad)** Phyllis is in just the same internal state as described above, and receives the same testimony from Tolloller and Mountararat, whom she has equal apparent reason to trust. But this time, although Mountararat is expressing knowledge he genuinely has, Tolloller is just making things up—he knows nothing in particular about which path Leila prefers. He's just trying to impress Phyllis.

In the Good case, Phyllis's reasons are exactly balanced between going to the left and going to the right—presumably then, she is warranted in taking either path. The Bad case is one in which Phyllis is internally just the same as she is in the Good case.[63] So if internalism is correct, then her reasons should be indifferent in the Bad case too. So given KR, it seems that Phyllis has *all the same* reasons she has in the good case to go to the left, and *strictly less* reason to go to the right.

The obvious difference between the two cases is that in the Good case, but not the Bad one, Phyllis knows that Leila said she usually takes the right path. In both cases she knows that Tolloller *said* that Leila said she usually takes the right path. This shared reason to go to the right seems exactly symmetrical with this shared reason to go to the left: Mountararat said that Celia usually takes the left path. But the other shared reason—that Celia usually takes the left path—seems in the bad case to have nothing to balance against it. So it looks as if, in the bad case, Phyllis's total reasons are asymmetric; they favour going left. So this looks like a challenge to the combination of KR and internalism (and so also to ER). And, insofar as there is a problem with the combination of internalism and KR, one might well think it is the latter that ought to go. For it is intuitively plausible that even in the bad case, Phyllis's reasons are neutral between the two paths.

[63] Although as I have suggested, I am rather open-minded about just which features—even external-world features, bear on how one is 'internally', I assume that whether some external party's testimony is true is *not* an internal difference. Only a rather extreme kind of disjunctivist would argue that there's no epistemically relevant state in common between gaining knowledge from good testimony and being misled by false testimony. To my knowledge, no one has defended a view of this kind in print, although Ram Neta has suggested to me some sympathy with it in conversation.

Note that a very closely related version of this challenge can be run in the realm of evidence and evidential probability; the most direct way to see this would be to consider just the same cases just mentioned, and consider Phyllis's evidential probability for the hypothesis that *she ought to go left*. In the bad case, it appears that more of her evidence favours this hypothesis than opposes it, for exactly the same reasons just rehearsed.[64] It is also not difficult to construct versions of this kind of case where one is not considering evidence with respect to normative propositions. (Consider, in the cases above, the evidential probability for Phyllis of the proposition that *there are more fairies to the left than to the right*.) This is not a direct challenge to the approach defended with respect to justified belief in Ch. 4, because the latter view did not relate justification so directly to evidential probability—the view is that one is justified just in case one is internally like someone who knows (and so someone who has evidential probability 1); external factors changing evidential probabilities does not imply any kind of externalism about justification. (The case of warranted action is more stark because one action's being slightly better supported justifies action in a way that one hypothesis's being slightly better supported does not justify belief.) Still, it is a somewhat surprising implication of E=K that evidential probabilities can change by way of purely external factors.

One might accept a degree of externalism in these cases, but argue against the internalist intuitions that make it unpalatable. One way to do this would be to observe that these kinds of cases—cases in which purely external factors make for a difference in which action is warranted, or which hypothesis is most likely, must be *marginal*. It takes a special kind of case, where the reasons are otherwise balanced just so, for external factors to be able to make this kind of difference. So perhaps internalism is only *mostly* true.

Another possible option would be to try to defend internalism by arguing that the details of the relevant expected value function does not validate the kind of 'reason weighing' model supposed by the argument. The idea would be that it is a mistake to assume that, just because when we list the reasons in favour of each path, there's an apparent asymmetry, that the expected value of taking the two paths isn't in fact equal. Perhaps there is a 'swamping' effect that renders the differential reason irrelevant.

But there is also a simpler kind strategy available to the internalist. An important part of the argument is the assumption that, in the bad case, Phyllis's knowledge is asymmetrical—she *knows* (from Mountararat's testimony) that Celia prefers the left path, but she *doesn't* know (though she thinks she does) that Fleta prefers the right path. But one needn't admit this. An internalist could defend ER by arguing that in this case, Phyllis doesn't know *either* asserted content—all she knows is that she's been *told* those two contents. Then the external asymmetry with respect to who's telling the truth would't extend to her evidential situation. Broadly speaking, I can see two ways an internalist might defend this kind of line.

The first is to argue that testimony is not in general a way to achieve knowledge. When one receives testimony that p, even from a trustworthy source who himself knows that p, one does *not* come to know that p. Perhaps one only gets to know that one's source *said* that p, and achieves reasonably high rational credence that p. This is, to be sure, a skeptical stance towards testimony, but it is not unprecedented—as Jennifer Nagel observes, it was John Locke's view.[65] If Phyllis doesn't

[64] It is no accident that the same considerations apply; indeed, if the Kearnes/Star view described in §5.11 is right, reasons just *are* evidence for oughts of this kind.

[65] See Nagel (2014, pp. 73–6), citing Locke (2008, §4.15). Nagel argues, against Locke, that this combining skepticism about testimony with non-skepticism about perception (as Locke does) is badly motivated, since perception is susceptible to error in the same way that testimony is. Whether this is right, however, seems to me to depend on the foundational epistemic questions discussed in Ch. 3; it seems not obviously badly motivated to suppose that disjunctivism about perception is true, but that disjunctivism about testimony isn't. (It is intuitively plausible that I *see for myself* that I have hands, but that I have to rely on trust in others to know who is President.)

know which path *either* fairy prefers, but only knows that someone told her something about each, the symmetry is restored. Still, this line is likely to strike most epistemologists as *too* skeptical; even if, being limited to testimony, it doesn't imply that we can't know *anything*, there is a great deal that we take ourselves to know that seems reliant on testimony. (To use one of Nagel's examples: if the Lockean view is right, then none of us know where we were born.) As is so often the case, however, this skeptical stance may be moderated to a considerable degree by contextualism; I'll elaborate that option further in the next section.

The second way of restoring symmetry in the bad case, consistent with KR, by denying testimonial knowledge of Celia's preference, is to argue that she can't gain knowledge by testimony *because of circumstances that obtain in this special instance*. Although one *can* gain knowledge from testimony under ordinary circumstances, one might argue that in cases like this one, circumstances fail to be ordinary in a way undermining of testimonial knowledge. The fact that Tolloller, who is in general no less trustworthy than Mountararat, is lying to Phyllis about a fairy's preferences, might be just the sort of thing that prevents the environment from being suitable for the receipt of testimonial knowledge. After all, from Phyllis's point of view, the method she is using for her two testimonial beliefs is *extremely* similar; it is not at all ad hoc to suppose that if the method is delivering a false belief in one case, it can't deliver knowledge in the other. One might hold this on general grounds of safety considerations,[66] or one might assimilate the case more specifically to a fake-barn-style Gettier case.[67] This is a version of a response Ram Neta gives to a challenge from Mark Johnston to disjunctivism about perception. Johnston's challenge is that it seems implausibly to be an implication of the latter that a subject can tell, in a case where he alternates between verdical perception and hallucination, which is which; Neta points out that the fact that the subject is prone to hallucination may well be preventing the perception from occurring and delivering knowledge. (Neta, 2008a, p. 323).

This move seems very plausible in this instance, but I am not sure whether it will be sufficiently general to vindicate ER and internalism. It is not essential to this kind of challenge that the reasons on each side be provided from the same kind of evidential source—any case in which the reasons are balanced before external changes undermine reasons on one side only will do. As remarked above, making the sources the same helps us to *recognise* these cases as such, but it's not a requirement for their *being* such. Still, maybe this kind of move can work in concert with the 'mostly internalist' fallback—the only counterexamples to internalism are cases that will be highly marginal and difficult to identify. Alternatively, one might try to extend the same story to cases even when they involve different evidential sources—maybe the fact that someone generally trustworthy is lying to Phyllis makes it harder for Phyllis to know *anything*, whether or not by testimony. (Maybe subjects in fake barn country can't know by sight whether an object is a silo, either.) Because I am unsure whether such a posit can be motivated on general grounds, I am unsure about the prospects for defending internalism purely along these lines.

Once we overlay our contextualist apparatus on knowledge and reason ascriptions, however, we may find a better motivation for this sort of internalism.

5.15 Contextualism and Symmetry

As mentioned in the previous section, contextualism about reason and knowledge ascriptions can help make the symmetry-restoring knowledge denials more plausible. In fact, Ram Neta invokes contextualism as one way of explaining the fake-barn-style move mentioned in response to Mark

[66] See Williamson (2000, Ch. 5), Sosa (1999), Lasonen-Aarnio (2008), and §2.9 of this book.

[67] Compare for instance Lewis (1996, p. 557): on Lewis's picture, by the Rule of Actuality, the possibility that Tolloller is lying is a relevant case; so, perhaps by the rule of Resemblance, the possibility that Mountararat is lying is also relevant.

Johnston's challenge to disjunctivism.[68] Neta's suggestion, adapted for the *Fairy Paths* case, is that even if the fact that one person is lying doesn't bear on whether the other person can transmit knowledge to Phyllis, the *discussion* of that possibility can move *us* into a conversational context relative to which 'Phyllis knows either content' expresses a falsehood. Mention of the possibility that someone is lying—whether or not that possibility obtains—moves one into a conversational context relative to which testimony cannot suffice for 'knowledge' of the content asserted.

One clear advantage of this version of the skeptical response is the same as the advantage for contextualist concessions to skepticism generally: one taking it needn't concede that ordinary ascriptions of testimonial knowledge are always false; in contexts in which the possibility of being misled is properly ignored, acceptance of testimony does suffice for satisfying 'knows'. But if, as in the context one was in when reading the fairy paths examples, one is considering such possibilities, neither testified content counts as 'known' by Phyllis. So, according to the metalinguistic generalization of KR under consideration, there is no challenge to the symmetry of Phyllis's reasons; so there may be no counterexample to internalism. Furthermore, this kind of solution *might* even extend to the kinds of cases where the symmetry is less transparent, due to differential sources of evidence. (Whether it definitively *would* do so depends on some more substantive and complex questions about the metasemantic rules governing the relevance of possibilities given contexts and indexes.)

One might worry that this solution is insufficiently general to be a full vindication of internalism. Consider, for example, the sentence, 'Phyllis's reasons are symmetrical between the two paths', uttered in a nonskeptical conversational context in which testimony counts as a source of 'knowledge', but evaluated relative to the bad case described above. One might worry that a contextualist line like the one we're looking at would imply that, since we're in a context where we're not worrying about the possibility of false testimony, Phyllis's attitude towards Mountararat's asserted content will amount to 'knowledge'; but even if we're not being skeptical about testimony, 'knowledge' must remain factive, and so Phyllis can't 'know' that which Tolloller told her. However, it's not clear how serious a problem this is. For one thing, any counterexamples to internalism that depended on this kind of mismatch between the speakers' assumptions and the subject's situation will be 'elusive'—any time one focused on these features of the situation, one would end up in a more skeptical context relative to which the symmetry would be restored. For another, it is very natural to invoke this kind of strategy *along with* supposing that a fake-barn-like situation makes it more difficult to get knowledge, *regardless* of whether this fact is recognised by any speakers. In the terminology of Ch. 1, the conversational context sets the 'epistemic standard' necessary for knowledge, but epistemic standards themselves can encode sensitivity to the epistemic suitability of the subject's environment. (Even relative to a given standard, it's harder for a subject to 'know' if that subject is in a fake barn situation. See §1.8.)

The contextualist framework may also give stronger prospects for generalising to cases of evidential symmetry involving different evidential sources on each side. It was a simplifying stipulation, in the Fairy Paths cases, that Phyllis's opposing reasons came from very similar sources—they were cases of testimony, with Phyllis epistemically indifferent between the sources, i.e., she rationally trusted them each equally. But, as mentioned in the previous section, one can construct a more gen-

[68] He does not commit to contextualism in this paper, but observes that there are two options along these lines: one, which calls the 'subject contextualist' option, says that the *subject's situation* is such that one can't get knowledge from the source in question—in particular, the subject is in a fake-barn-like Gettier situation. (As indicated in §1.2, I do not prefer the label 'contextualism' for this kind of view.) The other option, what Neta calls the 'attributor contextualist' option, has it that 'whether or not we can truthfully claim that Henry knows that *that is a barn* depends on features of our context (for example, whether or not we're thinking about [a fake barn case]'. Neta (2008a, p. 319) While Neta does not explicitly characterise this as a metasemantic discussion—it is highly debatable whether contextualists think that whether we can truthfully attribute knowledge depends on our context (see Sosa (2000, p. 4))—it does seem clear that he is talking about the view I'm describing as contextualism (see for example his p. 321, where is makes explicit that on this view, the contents of words are changing).

eral version of the worry. Suppose that Willis, like Phyllis, wants to find a fairy, and is wondering which path to take. Like Phyllis, he gets trustworthy testimony that Celia has gone to the left. But Willis's counterbalancing reason, unlike Phyllis's, is not based on testimony—it's based on something else that carries the same amount of epistemic weight. We can suppose, for instance, that it's based on induction. (Any non-basic form of knowledge will do.) He observed Leila going to the right on each of the past four days, say, and concludes on this basis that she went to the right again today. As a result, he has equal (rather strong) support for the proposition that *Celia went to the left* as he does for the proposition that *Leila went to the right*. (Our intuitions aren't precise enough to justify precise claims about just what kind of source would balance just how many days' track record, but it does seem clear that there should be some possible ways of spelling out the case such that things are perfectly balanced in this way, even if we wouldn't expect to be able to recognise them.) In the good case, relative to a nonskeptical context, Willis 'knows' that Celia went to the left, and also that Leila went to the right (and the testimonial and past evidence that puts him in a position to 'know'). But if one of those two sources is, unbeknownst to Willis, unreliable in this instance—if he is receiving false testimony, or if this is an unusual day for which past days are poor evidence—then it may seem that the symmetry can be broken by external considerations. On the contextualist version of the view, the problem comes up when one is in a conversational context that is more skeptical about one kind of source than the other. (Suppose for instance that one is entertaining the possibility that this is an unusual day, but not that Mountararat is lying.)

Depending on the details of one's metasemantic story about how conversational contexts fix epistemic standards, one may or may not think this kind of asymmetry is possible. On certain ways of thinking about epistemic standards, a kind symmetry may well be enforced. Consider for instance the approach of Keith DeRose, discussed in §1.8. Here again is DeRose's approach to epistemic standards:

> An important component of being in a strong epistemic position with respect to P is to have one's belief as to whether P is true match the fact of the matter as to whether P is true, not only in the actual world, but also at the worlds sufficiently close to the actual world. That is, one's belief should not only be true, but should be non-accidentally true, where this requires one's belief as to whether P is true to match the fact of the matter at nearby worlds. The further away one can get from the actual world, while still having it be the case that one's belief matches the fact at worlds that far away and closer, the stronger a position one is in with respect to P. DeRose (1995, p. 34)

If one thinks, as DeRose does, of epistemic standards as encoding a kind of 'modal distance' that, along with the index, defines the relevant set of possibilities, then increasing standards will diminish knowledge from non-basic sources together. In the case of Willis above, we can think of things this way: relative to an index in which Willis's testimonial judgment and inductive judgment each deliver knowledge, the nearest world where the testimony yields a false belief is the *same distance* as the nearest world where the inductive method yields a false generalisation. (This is a way of encoding the stipulation that the two evidential sources are *equally strong*.) So any change in conversational context that prevents one source from delivering 'knowledge' would likewise prevent the other. On this kind of approach, symmetry would be built in, and internalism may well be very plausible.

However, as I indicated in §1.8, the assumption that sets of relevant alternatives are settled in this kind of symmetry-enforcing way is an additional commitment, not motivated by the basic Lewisian framework. Prima facie, there seems to be no problem with pairs of epistemic standards, each of which is more skeptical in some respects, but less so in others. If, for instance, one can vary the degree to which testimony and induction count as 'knowledge'-producing *independently*, then there will have to be contexts in which asymmetries of this kind could arise. Does this mean that there

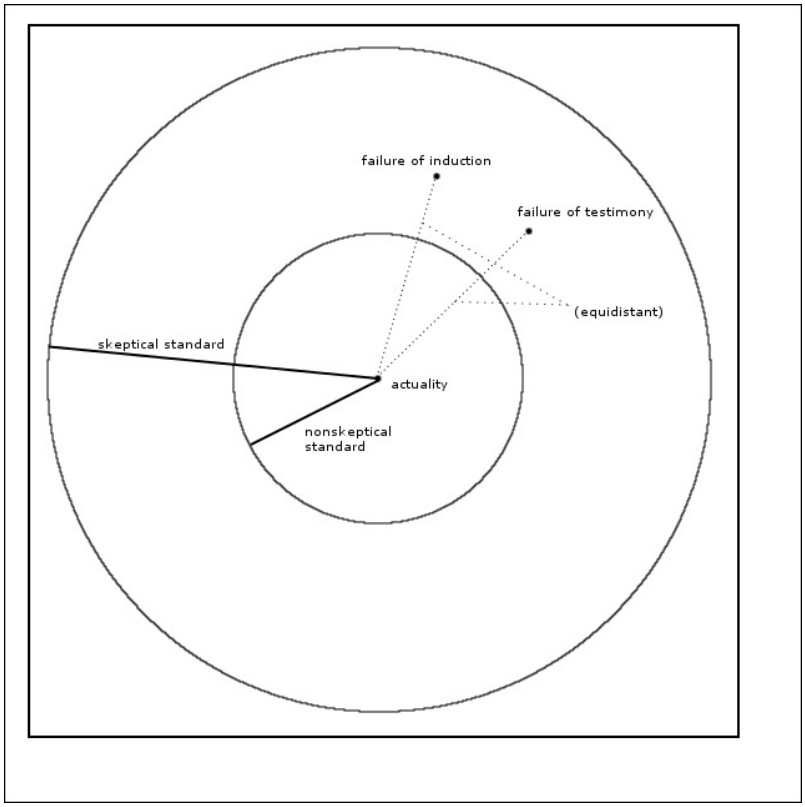

Figure 5.2: Any standard in which failures of one source are relevant is one in which failure of the other source are also relevant.

could be two contexts in which, in discussion of the same subject in the same situation, a sentence like 'Willis's total reasons best support going left' differ in their truth values? It does if we allow, as seemed intuitive, that the combined reasons on the one side—that Celia went to the left, and that Mountararat said that Celia went to the left—make for stronger reason than the single opposing reason that Tolloller said that Leila went to the right. Whether this is so depends on the details of the expected value function, or the probability function, described in §5.11. As I mentioned in the previous section, an internalist might reach for a 'swamping' explanation whereby Willis's extra reason in favour of going left makes no difference, given the reasons he already possesses. The challenge here will be to develop the story about swamping in a way that is plausibly generally motivated. I am unsure whether there are strong prospects for such a story. So I am inclined to think that, if one's metasemantic story about the fixing of relevant alternatives does not build in a kind of symmetry, a contextualist who embraces KR may end up committed to a denial of a certain kind of metasemantic statement of internalism. *Perhaps* this consideration should count as a mark in favour of versions of contextualism that include a kind of symmetry. But the result may not be as bad as it may seem. Suppose our contextualist admits that sentences like (13) can be true in some contexts and false in others:

(13) Willis's total reasons best support going left.

This implication enjoys a certain similarity to that posed in the objection to contextualism consid-

ered in §5.3, but the difference between it and that objection still leaves room for manoeuvre. The objection concerned 'ought' statements—one was concerned to recover the sense of substantive disagreement between people in distinct contexts who affirmed and denied a sentence like (14).

(14) Willis ought to go left.

One might think that conceding contextualism about R will force one to do the same about (14), on the grounds that one always ought to do what one's reasons best support. But any so straightforward argument would commit the contextualist fallacy. One can accept contextualism about (13) without contextualism about (14)[69] if one declines to endorse a *metalinguistically general* statement of such a link between 'reasons' and 'ought'. One can hold that one always ought to do what one's total reasons best support, but allow the existence of some conversational contexts relative to which that which one 'ought to do' differs from what one's 'total reasons best support'. (Obviously, such contexts cannot include *this* one I am presently writing in.) We may allow that there are *some* contexts where that which satisfies 'reasons' do not track facts about what one ought to. In particular, contexts that have the kind of asymmetrical structure described above may gerrymander 'reason' talk in a way that distorts it from the overall normative judgment. Such contexts, even if they are possible, are bound to be somewhat unusual; perhaps as a strong default, epistemic standards do not exhibit this kind of asymmetry. Notice that the kinds of cases that motivate the challenge to contextualism, like the case introduced in §5.3 of Strephon who embraced his Aunt, do not involve differential skepticism with respect to particular kinds of sources. We needn't even admit that sentences like 'S ought to Φ even though S's total reason weighs against Φ-ing' can ever be true, as explicit discussion of what one ought to do may well trigger conversational contexts in which 'reasons' are more relevant for what one ought to do.

5.16 Internalism and Basic Knowledge

What of internalism? As I have been using the term, the 'internal' comprises one's cognitive home—it is a matter of one's basic evidence and cognitive processing. (As throughout this book, I do not assume that basic evidence supervenes on one's brain state; it may well include the objects of factive perceptual states.) On my Lewisian form of contextualism, articulated in Ch. 1 and clarified in Ch. 3, the basic evidence will count as 'known' relative to *all* conversational contexts. In the most skeptical possible context—one working with an epistemic standard according to which *all* possibilities are relevant—only the basic evidence, and that which follows from it, will count as 'known', and as 'reasons'. Internalism about warranted action, then, will be vindicated if the totality of the basic evidence settles what action is warranted. (One might, but needn't, accept the stronger claim that the totality of the basic evidence *grounds* the facts about what one ought to do. Supervenience is enough for present purposes.) None of the challenges to internalist versions of KR we've considered provide any motivation against this claim. It has been a crucial feature of all of the cases discussed that one is relying on an *indirect* source of knowledge—a case of testimony, say, or a shopping list. These are not at all accidental features; the central move in constructing these arguments involved manipulating the knowledge without changing the internal. But in the special case of basic knowledge, that knowledge *is* internal. So we've seen no challenge to the idea that *basic* knowledge fully determines what one ought to do.

It is natural at this point to wonder about the degree to which the view ultimately sketched fits with KR and the broader knowledge-first project. Since I accept a distinction between knowledge

[69] Or, without contextualism *of the kind under consideration* about (14); contextualism about 'ought' claims is surely true for other reasons. For example, there are moral, prudential, and evidential 'oughts'.

and basic knowledge, and emphasise the importance of the latter, shouldn't this more properly be described as a defence of a *basic* knowledge first stance? And why accept the truth of KR—wouldn't the thesis that *basic* knowledge comprises one's reasons do just as well or better? There is a sense in which I agree with both of these thoughts. As highlighted in the Introduction, I think of the knowledge-first program as a claim about *relative*, rather than *absolute*, priority. One doesn't refute it by showing that there is something *else* that plays central roles too. I do think that basic knowledge has a very important role to play in epistemology—I have been highlighting respects in which it may play a more central role than knowledge generally does. As for 'BKR', the view that one's reasons are all and only those propositions one *basically* knows, I think there *is* a sense in which that view better carves at the epistemological joints. The reason I have endorsed KR instead of BKR, however, is that the latter just doesn't seem to do justice to our ordinary ways of talking about reasons. In contexts and situations where no one is or should worry about the veracity of a bit of testimony, it is extremely natural to describe someone's reason as the content she learned through testimony. BKR predicts that this is a false reason ascription (assuming that testimony is not a source of *basic* knowledge). This is, to my ear, simply too revisionary a semantic theory. The dialectic here is exactly the same as that between a contextualist and a skeptic. Since I am interested in capturing nonskeptical uses of the language as true, I opt for the contextualist treatment about reasons, just as I do about knowledge ascriptions themselves. But there is a sense in which, in the realm of practical reasoning, *basic* knowledge is where the action is.[70]

[70] Ichikawa (2016a) further explores the possibilities for the primacy of basic knowledge.

Chapter 6

Assertion

One of the roles knowledge is often thought to play is as a central—perhaps even a constitutive—*norm of assertion*. The idea is that knowledge plays a distinctive governing role in the practice of assertion: you should only assert things that you know.

There are many norms that govern acts of assertion: it's reasonably plausible, for instance, that you should also only assert things that won't incite violence, or that aren't terribly rude, or that don't contain racial pejoratives. Some epistemologists, such as Rachel McKinnon, have supposed that these kinds of considerations ought to be considered together with epistemic norms like the knowledge norm. McKinnon writes:

> There are a number of social conventions surrounding the appropriateness of some assertions. Although an assertion may be true, and even a case of knowledge, it may be impolite to assert during a department meeting that your boss has very bad breath. So there may be important conventional aspects to assertoric norms, *and this should be reflected in the articulation of the norms*.
>
> Additionally, there are important pragmatic aspects relevant to assertibility. There are many things that we do with assertions that may bear on assertibility conditions. For example, assertions are typically taken as reasons for action. Although there may be a picture of a fire on the wall of a crowded movie theater, it would be a rather terrible idea to suddenly loudly assert "There's a fire!" People will wrongly interpret this true assertion as a reason to stampede for the exists (naturally thinking that there's a real fire in the theater, rather than just a painting). Since this predictably bad outcome is a likely consequence of the assertion, it seems to make the assertion inappropriate. Understanding the pragmatic implications of assertion, including what conditions are required for licensing an interlocutor to take one's assertion and act on it, might therefore bear on whether one may or ought to assert. *And this, too, should be reflected in the articulation of the norms of assertion*. McKinnon (2015, p. 3, my emphasis)

But McKinnon may be moving a little hastily from her observation that conventional and pragmatic norms apply to assertions, to her conclusion that the norms *of assertion* should reflect them.[1] Certainly, defenders of the knowledge norm tend to think that knowledge has a *special* normative role to play in governing assertion, but that this is consistent with the clear fact that other norms are

[1] McKinnon herself may agree; later in the book, she clarifies that politeness norms, being external to assertion, are not norms *of* assertion. (McKinnon, 2015, p. 40). But she goes on to argue that other conventional and pragmatic norms *are* internal to assertion.

in place. Notice for instance that the norm to be polite, or to refrain from inciting violence, seems to be both *situation-specific* (there's a time and a place for being rude) and *derivative* from more general behavioural norms that are not specific to assertion (in a given situation, one should assert politely because one should, in that situation, be polite). By contrast, it may be that the knowledge norm governs assertion in a *constitutive* and/or *non-derivative* way. If it is constitutive, then part of what *makes* a given practice the practice of assertion, instead of some other practice, is that it is governed by a knowledge norm; if so, then it will always be in place for any assertion. If it is non-derivative, then it is not simply an instance of a more general norm, applied to the case of assertion. I am sympathetic to these stronger interpretations of a knowledge norm of assertion. Matthew Benton (2014) argues very effectively that the knowledge norm is a nice way of giving flesh to Grice's 'Maxim of Quality', and that so understood, it is both a constitutive and a non-derivative norm for assertion. (The Griceian system affords a special significance to the Maxim of Quality; Grice's other rules derive from more general norms.) Call the thesis that knowledge plays this special normative role in the realm of assertion 'KA':

KA. Knowledge is a constitutive, nonderivative norm of assertion.

Some epistemologists have argued that positing a constitutive epistemic norm along the lines of KA is necessary to explain the epistemology of testimony.[2] A case might well be made for the stronger claim that knowledge is the one and only such norm: that any other norm governing assertion would be either nonconstitutive or derivative (either from the knowledge norm, as in the case for the norm that one should only assert what one believes, or from some other more general norm, as in the case for the norm that one should assert only what won't cause people to panic unnecessarily). One might also hold the converse uniqueness claim—that *assertion* is the distinctive illocutionary practice governed by a knowledge norm.[3] While I find these strengthenings plausible, the weaker version I've articulated as KA is controversial enough, and adequate for the purposes of this chapter, so it's the main claim I'll be considering.

As with other knowledge norms, as well as other theoretical principles involving knowledge, a contextualist has various choices for how to interpret knowledge norms. Broadly speaking, the project of this chapter will be to examine this choice point from several angles.

My first central aim in this chapter will be to rebut arguments connecting contextualism to KA. In §§6.2–6.3, I'll examine the so-called 'factivity problem' for contextualism and KA; in §§6.4–6.5, I'll consider Keith DeRose's argument that KA provides a straightforward motivation in *favour* of contextualism. Both kinds of arguments, I think, fail in an instructive way. I'll draw some broader methodological morals in §§6.6–6.9, before turning to the positive section of the chapter. In §§6.10–6.14, I'll develop what I think is an underexplored proposal for how a contextualist can implement something in the neighbourhood of KA.

As a preliminary to these projects, however, I'd like to begin in §6.1 by examining a challenge and alternative to KA that Jason Stanley has pressed; the contextualist framework I've been developing allows one to develop Stanley's insights in different—and to my mind, rather interesting—ways.

[2] See Goldberg (2015), who does not endorse KA, but who is a clear proponent of the view that *some* constitutive epistemic norm is essential. Casey Rebecca Johnson argues against this line of thought, suggesting that we can explain knowledge via testimony and other related phenomena by reference to contingent social conventions, but even she allows that there might be other reasons to adopt such a constitutive norm—see Johnson (2015, p. 371).

[3] Williamson endorses both distinctiveness claims. (Williamson, 2000, p. 241). See McKinnon (2015, pp. 44–9) and Pagin (2015, pp. 7–8) for helpful clarifications and discussion.

6.1 Stanley and the Certainty Norm

Jason Stanley (2008) observes that some of the data often claimed in favour of the knowledge norm of assertion does not distinguish between KA and the *certainty* norm, CA:

> CA Assert that *p* only if you are certain that *p*. (Stanley, 2008, p. 48)

One of the canonical motivations for KA is that it is well-positioned to explain the infelicity of the Moore-paradoxical '*p* but I don't know that *p*'.[4] The second conjunct entails that the first conjunct is, according to KA, unassertable. So any such assertion would be an assertion of the fact that one is unlicensed to make it. Stanley thinks the problem for this story is that, Peter Unger's remarks to the contrary, it is insufficiently general. For first-personal denials of *certainty* also generate Moore-paradoxicality: '*p* but I am not certain that *p*' is unacceptable in the same way that '*p* but I don't know that *p*' is. As Stanley (p. 47) observes, Peter Unger, who defends the knowledge norm in part on the grounds of its explanation of Moore-paradoxicality, is well aware of this symmetry. (Unger, 1975) Unlike Stanley, however, Unger thinks that the knowledge norm is able to explain the unacceptability of *both* knowledge-denying and certainty-denying Moore paradoxes, because on Unger's view, knowledge *entails* certainty. Stanley denies that knowledge requires certainty, citing sentences like 'she knows that *p*, though she's only somewhat certain of it'. Stanley (2008, p. 46) But, Stanley argues, a *certainty* norm of assertion can explain all of the relevant data. It explains the certainty-denying Moore paradoxes in the familiar way, and it explains the knowledge-denying version because the relevant certainty required for assertion must put one in a position for knowledge.[5]

[4] Unger (1975, pp. 258–60), Williamson (2000, pp. 253–4), DeRose (2002, p. 181).

[5] In fact, I doubt whether Stanley succeeds in capturing the data motivating the knowledge norm for the certainty norm. He would have a relatively easy job if CA were *strictly stronger* than KA. But it isn't. The relevant notion of certainty here is epistemic certainty, according to which 'one is certain of a proposition p if and only if one knows that p (or is in a position to know that p) on the basis of evidence that gives one the highest degree of justification for one's belief that p'. Stanley (2008, p. 35) Notably, S's having certainty that *p* does *not*, on Stanley's view, entail that S knows that *p*; it entails only that S knows *or is in a position* to know that *p* on a suitable basis. Here is Stanley's attempt to explain a derivative knowledge norm from a more basic certainty norm:

> Consider the proposition that there are no large Jewish elephants in my bedroom. This may have been an epistemic certainty for me five minutes ago, even though I did not know that there were no large Jewish elephants in my bedroom. I did not know that there were no large Jewish elephants in my bedroom, because I did not believe it, and I did not believe it simply because it didn't occur to me ever to entertain that possibility. Nevertheless, in this case, if I had entertained the proposition that there are no large Jewish elephants in my bedroom, I would have known it. The reason this counterfactual is true is because it is an epistemic certainty for me that there are no large Jewish elephants in my bedroom. So the fact that a proposition is an epistemic certainty for a person does not entail that the person knows that proposition. If a proposition is an epistemic certainty for a person at a time, then it does follow that the person is in a *position to know* that proposition. Being in a position to know a proposition is to be disposed to acquire the knowledge that the proposition is true, when one entertains it on the right evidential basis. Since epistemic certainty entails possession of this dispositional property, utterances of [knowledge-denying Moore paradoxes] are odd. (Stanley, 2008, pp. 48–9)

But since certainty does not entail knowledge, on Stanley's view, it is not at all clear how a certainty norm could explain the general infelicity of '*p* but I don't know that *p*'. One can't *know* both conjuncts, but it's unclear why one couldn't have *certainty* in both conjuncts. Stanley's thought seems to be that if something is certain, then *if* one asserts it, one must know it. But it doesn't follow from his explanations of these notions that this must be so—something might be asserted without being entertained on the right evidential basis. Suppose for instance that *p* is certain for me, but I don't know that *p* because I am ignoring my overwhelming evidence for *p*, and basing my belief that *p* on some bad evidence. Stanley does not appear to have offered an explanation for why I couldn't have certainty, in such a case, that I don't know that *p*. Indeed, it seems plausible that I might—my evidence, which includes some sort of introspective access to the source of my belief—might overwhelmingly establish that I don't know that *p*. But if so, the certainty norm predicts that '*p* but I don't know that *p*' should be assertable.

Given a contextualist framework for knowledge ascriptions like the one I've been suggesting in this book, there is room to resist Stanley's argument that knowledge does not require certainty. (Stanley too recognises that contextualism represents a way to resist his argument, although he takes himself to have ruled contextualism out.[6] The imperfect but perhaps tolerable 'she knows that *p*, though she's only somewhat certain of it' is assimilated to 'she knows that *p*, though she doesn't *know* that *p*,' where the emphasis in the latter case signals a shift to a more skeptical conversational context. Just the same diagnosis is given here as for quasi-tolerable statements of fallibilism as in 'she knows that *p*, even though there are some not-p cases that she can't rule out'.

Note also that there are strong motivations for supposing knowledge *does* require certainty. For example, I cannot hear any natural reading of a sentence like 'she is uncertain about *p*, but she knows *p*.' Even Stanley's own case sounds much worse with the order reversed: 'she's only somewhat certain that *p*, but she does know *p*'. In the same way, I don't find 'there are some not-*p* cases she can't rule out, but she does know that *p*' even quasi-tolerable. A contextualist can explain the asymmetry here by reference to the fact that it seems easier on the whole to raise standards than to lower them. ('I know that I'll teach epistemology next term, though I don't know that I won't die tomorrow' sounds bad; 'I don't know that I won't die tomorrow, but I do know that I'll teach epistemology next term' sounds worse.) See the parallel discussion in §2.3.[7]

6.2 The Factivity Challenge

Several authors have recently argued that there is a particular sort of problem with combining contextualism with the knowledge norm of assertion. The general idea is that there is a kind of self-defeat with respect to the contextualist's commitment to contextualism. (A different kind of alleged problem about combining contextualism with knowledge norms, associated particularly with John Hawthorne, and less specific to assertion norms as such, has already been mentioned a few times throughout this book—for instance, in §5.3. I'll discuss it with particular reference to assertion in §6.6 below.) The critique I have in mind has been pressed many times over the past decade, for instance by Crispin Wright (2005), Elke Brendel (2005), Peter Baumann (2008) and (2010), Daniele Sgaravatti (2013), Christoph Jäger (2012), Chris Buford (2009), and Anthony Brueckner and Chris Buford (2009). These critiques are all at their core the same: they challenge the idea that contextualists can know, or state, contextualism or that which follows from it.[8] I'll work most specifically with Buford's (2009) formulation, drawing connections elsewhere as appropriate.

Buford begins his discussion by observing quite rightly that it is *not* an implication of contextualism that one stops knowing things when the conversational standards raise. To suppose otherwise is flagrantly to commit the contextualist fallacy. So it would *not* be a good objection to contextualism to accuse it of committing to the truth of the last sentence in a dialogue like this one:

> Edith: I can't make out the flag on that ship; do you know what kind of men are aboard?
>
> Kate: Yes, I see a skull and crossbones flag. Those are pirates!
>
> Edith: But do you have any evidence ruling out the possibility that they're just *pretending* to be pirates? Have you seen them plunder any ships?

[6] Stanley (2008, p. 49).

[7] See also David Lewis's remarks on the general asymmetry of strengthening vs. weakening standards (not only in the case of 'knows'). Lewis (1979b, pp. 352–3, 355)

[8] Brueckner and Buford (2009) is cast not as an objection but as a defence of contextualism, but they, like the other authors listed above, do attribute to contextualism a kind of inability to stand up for itself. I agree with Baumann (2010, p. 87) that this would be a drawback—I also think it's one that derives from many of the same problematic assumptions that underwrite the more aggressive stances. So I categorise them with the critics for the purposes of this section.

> Kate: No, I was just going by the flag. I can't tell whether they're pretending.
>
> Edith: So you admit that you didn't know that they're pirates.
>
> Kate: No, I did know that they're pirates; it's just that, now that we're talking about the possibility that they're pretending, I don't know that any more.[9]

For reasons that should be clear by this point in this book, contextualism does not imply that Kate's final utterance here is correct. Assuming that when Edith raised her skeptical possibility, they shifted into a more skeptical context relative to which Kate does not 'know' that they're pirates, 'know' now picks out a more demanding relation. So Kate should not have said that she 'did know' but doesn't any more; relative to her high standard, she *never* 'knew' (so 'I never knew' would be true if she said it), even though her previous knowledge ascription was true.

However, Buford argues, a modification of this confused argument against contextualism, combined with the knowledge norm of assertion, *does* result in a genuine problem for contextualism. They key is to replace Kate's problematic final utterance with something like this:

> Kate: No, I don't know now, and I didn't know a moment ago, that they're pirates. But I stand by what I said before—it was true.

I should make two observations at this point. First, Buford's objection is *not* that Kate's modified final claim here is fine by contextualist lights, even though it is straightforwardly intuitively problematic. Arguably, it is an intuitively problematic utterance, and this *is* an influential objection to contextualism—data along these lines are among the primary motivations for *relativism* about knowledge ascriptions.[10] A comparative investigation into the merits of contextualism versus relativism is beyond the scope of this book, but the reader won't be surprised to learn that I'm not convinced by this objection.[11] At any rate, Buford's concern is different; he argues that the appropriateness of this utterance raises a particular problem for contextualists who endorse the knowledge norm of assertion—I'll examine this thought in some detail below.

Second, unfortunately, Buford is not specific about just what the felicitous utterance that generates the challenge is. He never explicitly lays out the alternate version of the dialogue he has in mind. He does however quote a passage from Keith DeRose pointing to a claim somewhat like this one; so this is my best rational reconstruction of the argument Buford has in mind.[12]

Assume that, as a contextualist might well think would happen in a case like this, Kate and Edith are in a relatively non-skeptical context C_L at the start of their discussion, but that they shift to a more skeptical context C_H when Kate mentions the possibility that the sailors' flag might not indicate their true identities.

Buford's complaint against the contextualist is that Kate's final assertion in C_H of 'I stand by what I said before—it was true' (where what she said before is what she said with her utterance, in C_L, of 'yes' in response to 'do you know what kind of men are on board'), though *true* by contextualist

[9] This dialogue is my own, adapted from one Buford gives to make the point on Buford (2009, p. 113) (citing Yourgrau (1983, p. 183).

[10] See MacFarlane (2005), MacFarlane (2007), Brogaard (2008), Richard (2004), Khoo (2017).

[11] One (small) part of the reason I'm unconvinced by this objection is that, as I'll lay out below, I think it's a mistake to suppose that contextualists are committed to thinking this is a fine thing for Kate to say, although I do admit that some of the related relativist-friendly data are not undermined by this move. See McKenna (2014), Carter (2013), Dowell (2011), Huvenes (2014), or Huvenes (2017) for a sense of some of the standard ways in which contextualists can respond to this kind of worry.

[12] Jäger's corresponding commitment is that a contextualist can appropriately assert the truth of contextualism, even in high-standards contexts. (Jäger, 2012, p. 492)

lights, should not be countenanced as *assertable*.[13] Here is Buford's line of thought, focused on Kate's C_H assertion to the effect that her C_L self-ascription of knowledge was true: if the knowledge norm of assertion is correct, then anything assertable must be known. So when Kate says, in C_H, that her assertion in C_L (that she 'knew' they were pirates) was true, she must *know*, at that time, that her assertion in C_L was true. Buford supposes that, for the purpose of the utterance in C_H, the relevant standard here is that provided by C_H. (He seems to be assuming a particular subject-centered metalinguistic generalization of the knowledge norm.) If so, then for her last utterance to be assertable, Kate must know by the higher standards of C_H—she must know_H—that she previously knew by the lower standards of C_L—she knew_L. The difficulty is that if she knows_H that she previously knew_L that they're pirates, it seems, contrary to the contextualist treatment of skepticism, that she ought at that time to know_H—not merely to know_L—that they're pirates.

As Buford explains it, this is because of a general closure principle on knowledge_H. If someone knows_H that p, and knows_H that p entails q, then one is in a position to know_H q. (As Buford correctly notes, maintaining consistency with such closure principles has been a historical motivation for contextualism.) But on the plausible assumption that Kate knows_H that knowledge_L is factive—that anything anyone knows_L is true—then Kate's knowledge_H that she used to know_L that they're pirates would put her in a position to know_H that they're pirates.[14]

I can see no grounds for supposing that Kate wouldn't know_H that knowledge_H is factive. The difference between C_L and C_H is that the possibility that the flag might not reflect the true identity has been raised. Although this has a skeptical effect on the conversational score, there's no reason to think it should introduce 'possibilities' as relevant in which subjects know falsehoods. This is only a relatively moderate skeptical scenario.[15] So, assuming closure of knowledge_H, Kate's knowledge_H of the factivity of knowledge_L, and a subject-centred metalinguistic interpretation of the knowledge

[13] Perhaps a more canonical way of setting out this dialectic would make a dialogue with simpler and more explicit knowledge claims and references to them. Readers familiar with the literature in this area might be more used to seeing a dialogue like this one:

> Edith: Can you see what kind of men are aboard?
>
> Kate: Yes, I see a skull and crossbones flag. I know that they are pirates!
>
> Edith: But do you have any evidence ruling out the possibility that they're just *pretending* to be pirates? Have you seen them plunder any ships?
>
> Kate: No, I was just going by the flag. I don't know whether they're pretending. Still, what I said earlier is true.

I find it extremely difficult to have useful semantic intuitions about dialogues like these; they strike me as very stilted. Ordinary speakers almost never make bare 'I know that p' assertions in nonskeptical contexts in which p is the central piece of information being contributed. (One might say 'I know that p' if one is working hard to convince a reluctant interlocutor to believing one—trust me, I'm an expert, I *know* it—or if one is cutting off an unnecessary explanation—you're preaching to the choir; I already know that p.) I don't think I've *ever* looked into the distance, seen that p, and then announced to my conversational partners, 'I know that p'. So I've made some small effort above to make the dialogue a bit more natural; the cost is that the commitments about knowledge are less explicit. (The central assertion that commits Kate to satisfying the lower 'knows' standard is her assertion 'Yes' in response to a knowledge question.) No doubt, still more natural versions of these dialogues are possible and in some respects preferable, though they may require yet further interpretation along these lines.

[14] Although there is a temporal change over the course of the conversation, we may assume that the possibility that the gentlemen in question have ceased to be pirates *in the past few seconds* is irrelevant or ruled out. If you wish for some reason to be *very* precise about this, focus on Kate's epistemic state vis-à-vis the proposition that they're pirates at the particular time of utterance of the C_L assertion.

[15] Indeed, it's very much debatable whether it is even possible for there to be conversational contexts skeptical enough to allow as relevant 'possibilities' like the possibility that knowledge isn't factive (or that $2 + 4 = 9$, or that not all policemen are policemen, or various other 'rational impossibilities'). As I indicated in §1.7, I think that Lewis was mistaken to limit his 'possibilities' to metaphysically possible centred worlds, and so to insist that one's evidence might fail to rule out every possibility in which Hesperus is not Phosphorus. But I am not sure that the same goes for rational necessities. These—I have in mind the propositions that people sometimes call 'analyticities' or 'conceptual truths'—may well be such that there is *no*

norm of assertion, the propriety of Kate's last sentence in the dialogue above entails, contrary to the contextualist treatment of the case, that Kate knows$_H$ that they're pirates.

6.3 High-Standards Assertability of Low-Standards Knowledge

The key assumption in the argument is that Kate's utterance about her previous knowledge ascription is not merely true, but *assertable* in the high standards context. Buford's argument can't get off the ground without such assertability; if the fact that Kate knew$_L$ that they're pirates is simply one of the many truths that she isn't, in C$_H$, warranted in asserting, the contextualist can explain everything that's going on just fine. So, why should we suppose that such an assertion would be warranted? Buford says surprisingly little on this central question. He alleges that a meta-linguistic claim like 'I stand by what I said before—it was true' 'seems perfectly appropriate'. (Buford, 2009, p. 116)[16] I do not share in this 'feeling of appropriateness'. And for reasons very much like the ones Buford has articulated, a contextualist interested in the norm of assertion faces considerable pressure *not* to suppose that such claims are assertable. Indeed, as alluded to above, the *argument from disagreement* against contextualism takes as its starting-point that such claims *aren't* assertable.[17] So it does not seem to me that the contextualist faces pressure to say, as Buford says he should, that in the high-standards context one should be able warrantedly to assert that what one said in the low-standards context was true.

Another argument Buford offers here in favour of the contextualist's commitment to the assertability of such metasemantic claims, is that

> [a]s noted earlier, Contextualism is often advertised as a way of countering the external world skeptic. If the Contextualist cannot make meta-linguistic claims akin to [those at issue] . . . , then the Contextualist begins to look much like the reliabilist who asserts that *if* there is an external world, and our beliefs concerning it are formed as the result of a reliable belief-forming mechanism, then we have knowledge of the external world. Many see this as an unsatisfying response to the external world skeptic. (Buford, 2009, p. 116)

possibility that can be relevant in *any* conversational context in which they don't obtain. If so, a subject's basic evidence will *always* rule out any possibility in which they're false, for the vacuous reason that there are no such possibilities. Or, effectively equivalently, even if such 'possibilities' can be entered as relevant in conversational contexts, they are trivially *eliminated* by every possible body of evidence. (This does not imply that everyone always knows every rational necessity, because of the other differences between Lewis's view and my own version of Lewisian contextualism—see again §1.7.) Rational necessities, in the framework of Ichikawa and Jarvis (2013), are such that everyone always has conclusive reason to accept them. Any failure to accept them has to do with failure to *process* one's evidence, not with failure to have *enough* of it.

[16] Again, exegesis here is non-trivial, as Buford doesn't tell us *exactly* what meta-linguistic claim he's talking about, but it seems to me it must be one like the one mentioned above. Rather surprisingly, Buford suggests that DeRose himself considered such a claim to be assertable. This seems to be on the basis of this remark from DeRose: 'B cannot truly say, "I did know then that it was a zebra"; that would be like my saying, "I was here." B *can* say, "My previous knowledge claim was true," just as I can say, "My previous location claim was true." Or so I believe. But saying these things would have a point only if one were interested in the truth-value of the earlier claim, rather than in the question of whether in the *present* contextually determined sense one knew and knows, or didn't and doesn't.' DeRose (1992, p. 925) Against Buford, I do not read any claim about warranted assertability in this passage in DeRose; in contrasting the two utterances, he was, it seems to me, saying that one would be false and the other wouldn't, not that one would be false and the other would be warrantedly assertable. In effect, this is a discussion about how to resolve the context-sensitivity of the modal 'can'. Buford assumes that DeRose is using it to mean something like 'consistent with the norms of assertion'; I think it just meant 'consistent with saying truths'.

[17] See again e.g. MacFarlane (2005), MacFarlane (2007), Brogaard (2008), Richard (2004), or Khoo (2017).

Here are two observations in response.[18] First, although Buford is plausibly right that contextualism is 'often advertised' as a way of dealing with the skeptic, this is not, as I've emphasised in Ch. 3, a mandatory feature of contextualism, or its only motivation.[19] Second, and perhaps more to the point, Buford seems in this passage to be assuming that 'the Contextualist' is in a high-standards context relative to which he does not 'know' that radical skeptical scenarios don't obtain. This is a surprisingly widespread assumption. Christoph Jäger makes it in the context of a similar argument: 'Consider a conversational context in which the contextualist states his theory. Such contexts are paradigmatic epistemology or "philosophy classroom" contexts in which sceptical hypotheses are salient and taken seriously.'[20] Although it seems to be a widespread assumption that contextualists are in high-standards contexts, I see no reason to suppose that this must be—or typically is—so.[21] Although we contextualists are perhaps *somewhat* more likely to enter into very skeptical conversational contexts than ordinary people are—I'm happy to admit it likely that we think seriously and with an open mind about lottery paradoxes and brains in vats more than your average stockbroker does—this quirk of ours should not be exaggerated. We contextualists, just like everybody else, spend *most* of our time—including most of the time we spend engaged in conversations and writings about contextualism—in conversational contexts that do not take radical skeptical scenarios seriously. As I'm sitting here writing this book on contextualism, for example, I am not taking seriously the possibility that the world will end tomorrow, or that it came into being yesterday. I know where my dog and my wife are, even though at the moment the former is out of eyesight and the latter is out of the country. I'm writing in a context that generates an epistemic standard that is pretty typical of academic writing. If this book were written in a very skeptical context, the subject-centred knowledge norm would predict that I can't make just about any of the assertions of this book. I don't know relative to *very high* standards, for example, that Christopher Buford offered the argument quoted above. (I've downloaded a pdf file from a journal website, *apparently* a faithful copy of an article *apparently* written by Buford (a person who *apparently* exists), that *seemed* to contain words expressing such an argument…) But those aren't standards that are governing me now, and so I can make many assertions about things like what authors have published what arguments. This, even though I am *at this very minute* writing a book about contextualism and knowledge and epistemology. Although David Lewis held that discussion of skeptical scenarios leads to skeptical contexts ('that is how epistemology destroys knowledge'), that was a mistake.[22]

There is also a deeper problem for the assumption that metalinguistic utterances like Kate's must be assertable along these lines. Contextualism is a controversial semantic theory. Most people have never even heard of it; if Kate is among them, there's no reason whatsoever to predict that she'd offer the contextualist diagnosis of the dialogue. Even if we restrict our attention to those select few of us who know what contextualism is, many are confident that it's false, and many more don't

[18] The two observations in the main text undercut the analogy with reliabilism, but it is also dubious that reliabilism can do no better than such a conditional claim; reliabilists can categorically endorse knowledge of the external world, because they have reliable methods—those of psychology, for instance—that deliver the verdict that there is knowledge of the external world. The externalism in question applies at the higher-order level too. See Kornblith (2004).

[19] Keith DeRose also emphasises this point in DeRose (2009).

[20] (Jäger, 2012, p. 491). Brendel (2005, pp. 50–1) and Sgaravatti (2013, p. 111) offer arguments to this effect, based on the idea that engagement with skeptics inevitably raise epistemic standards arbitrarily highly. But I deny this latter assumption; see §1.7.

[21] For a point of agreement on slightly different grounds, see Blome-Tillmann (2014, pp. 51–3).

[22] (Lewis, 1996, p. 560). See §3.6, fn. 19 for why the 'destroys' language has to be non-literal. Note that even Lewis isn't quite as categorical on this point as it might appear: he holds that the *discussion of skeptical scenarios* will inevitably lead to skeptical contexts; but not all epistemology involves the discussion of such scenarios. (Note the antecedent of his p. 560 claim that '[u]nless this investigation of ours [discussing skeptical scenarios] was an altogether atypical sample of epistemology, it will be inevitable that epistemology must destroy knowledge.' But there *are* many respects in which fixation on skeptical scenarios is, if not quite *atypical*, at least inessential to a great deal of epistemology, including contextualist epistemology.

know whether it's true. In fact, in my view, even most—perhaps all—*contextualists* don't know that contextualism is true. (As I've been doing throughout this book, I'm writing in a context with moderate, rather than highly skeptical, epistemic standards.)[23]

Contextualism is a theory that I accept, but I don't think the acceptance at issue here is a kind of outright belief (more on this in Ch. 7), and I do not assert its truth. It's just a mistake to suppose that, by the contextualist's lights, contextualism itself, or its controversial commitments, should be assertable in lax conversational settings, or that they're known, even relative to low-standards contexts. I don't think anybody should go around asserting that contextualism is true. And *I'm* both a professional epistemologist and a contextualist. *Kate*—the character that Buford and Jäger think should be asserting that her past utterance was true on the grounds that it was made in a low-standards context—has never even *heard* of contextualism. So I certainly don't think it's warranted to suppose that *she* should be able to assert consequences of contextualism in her high-standards context (or even in a low-standards one).

Finally, even if one *were* committed to Kate's knowing$_H$ that contextualism is true, *even that* wouldn't commit Kate to knowledge$_H$ that she knew$_L$ that the gentlemen in question were pirates, for the straightforward reason that contextualism is a semantic thesis about knowledge ascriptions, and accordingly has no implications whatsoever about whether any particular individuals are pirates.[24] Contextualism obviously also does not imply that *any* ascription of the form 'Kate knows that these men are pirates' is true. Since contextualism is consistent with the possibility that they're not pirates, it is consistent with the possibility that an ascription in *any* context of 'Kate knows they're pirates' is false. So even if Kate *did* know$_H$ that contextualism is true, that just wouldn't imply that she knows$_H$ the truth of any particular low-standards knowledge ascription.[25] Perhaps someone might argue—I don't find this particularly plausible, but I can certainly imagine an argument along these lines—that Kate has better epistemic access to the truth of contextualism than she does to the profession of the men on the ship. Perhaps the semantic fact is 'internal' in a way that makes it better-resistant to skeptical scenarios. Then maybe she knows$_H$ that contextualism is true, but doesn't know$_H$ that the men are pirates. Then she certainly won't know$_H$ that she knows$_L$ that they're pirates. The closest she'd get, given the relevant closure principles, is that she'd know$_H$ that, were she to utter 'I know that they're pirates' in C$_L$, she would be saying something relatively weak that *might*, for all she knows$_H$, be true. There's just no reason a contextualist need suppose that she'd know$_H$ that she'd know$_L$ that they're pirates.[26]

[23] Sandford Goldberg (2015) thinks that this and related facts cause trouble for the knowledge norm of assertion, since philosophers regularly assert such controversial claims. I am not nearly so confident as Goldberg is that philosophers often make outright assertions of such contested claims, or that it would be such a great cost to suppose that if they did, they'd be unwarranted. Much philosophy is written and spoken in a *speculative*, rather than an *assertive*, mode—and qualifiers like 'it seems to me that' are often used to weaken our commitments. (One might express this weakening by supposing that we're asserting a weaker content, like a content about how things seem, or one might consider these kinds of 'hedges' to signal weaker illocutionary acts than assertions. See Simons (2007).) Goldberg recognises this kind of strategy in response to his argument— Goldberg (2015, p. 249)—but I think he overestimates the commitments of the strategy—he seems at times to be engaging with the idea that philosophers *never* make controversial assertions, rather than the more modest claims that (a) they only *rarely* do so, and (b) when they do, it is unwarranted). I think that Goldberg also overestimates the assertoric force of philosophical writing. For example, although it is possible that I have inadvertently written in a more zealous way than I ought to have done, I am reasonably confident that at no point in this book, or in any of my other published writing, do I assert the truth of contextualism.

[24] Compare Freitag (2011) for a similar argument.

[25] So I disagree with Jäger, who writes that someone pursuing the kind of move I'm suggesting 'would have to argue that the contextualist can legitimately deny … that he knows when he asserts his theory, in [a high standards context], that there are low-standards contexts in which (it is true to say that) he knows that he has hands. The claim that there are such quotidian contexts, however, is a cornerstone of classical, anti-sceptical forms of contextualism.' Jäger (2012, p. 496) Since it is not a cornerstone of contextualism that I or anyone else has hands, it can't be a cornerstone *of contextualism* that low-standards 'knowledge' ascriptions about having hands are true.

[26] Note that in Brendel's version of the argument, the 'main thesis' of contextualism is first given as such a possibility

I conclude that contextualism faces no particular challenge deriving from the unassertability, in high-standards contexts, of contextualism or its commitments.[27]

6.4 DeRose 2002

In his 'Assertion, Knowledge, and Context', Keith DeRose attempted to motivate contextualism from the knowledge norm of assertion. This was the key line of thought:

> If the standards for when one is in a position to warrantedly assert that P are the same as those that constitute a truth condition for 'I know that P', then if the former vary with context, so do the latter. In short: The knowledge account of assertion together with the context sensitivity of assertability yields contextualism about knowledge. DeRose (2002, p. 187)

Put into a broad form, his argument ran thus:

1. Knowledge is the epistemic norm of assertion;
2. Assertability (even limited to epistemic considerations) varies according to conversational context; therefore
3. Contextualism about knowledge ascriptions is correct.

The point of the parenthetical in premise (2) is to recognise that there are irrelevant respects in which assertability varies according to context—for example, whether it is proper all things considered to assert something depends in part on whether it'd be rude to do so, and in different contexts, different assertions would be rude. Premise (2) says that, in different contexts, one can be licensed *purely from an epistemic point of view*, whether or not one is *all things considered* licensed, to assert differentially. For example, in a context in which a great deal hangs on whether p is true, one might not be justified in asserting that p even though the very same epistemic status might be good enough to warrant asserting that p in a context in which less is at stake.

Even with this clarification in hand, DeRose's argument is fallacious. In fact, it is doubly equivocal—the link between each premise and the conclusion is problematic. First, premise (1) is a premise about knowledge; to infer contextualism—a claim about the word 'knows'—is an instance of the contextualist fallacy. (Remember, knowledge ascriptions are sentences using the word 'knows'; premise (3) is equivalent to the claim that contextualism about 'knows' is correct.) So any argument like the one laid out in (1)–(3) equivocates on use and mention of 'knows'.

---

claim. Brendel (2005, p. 40) But later in the paper, her complaint against contextualism is that one can't assert, in the high-standards context, the *truth* of the claim about the low-standards utterance. (Brendel, 2005, pp. 49–50)

[27] Buford offers one more argument in his paper against the suggestion that contextualists shouldn't admit that the metasemantic claims in question are assertable: he finds implausible the apparent implication that, in a case like Kate's above, Kate would not be in a position to 'say anything about' her earlier assertion. By 'say anything about', Buford apparently means 'assert that it is true or that it is false'. (There's not even a prima facie challenge to the idea that Kate might say that she thinks it's true, or that it's probably true, or that she doesn't know whether it's true, or to the idea that she might perform any particular non-assertive illocutionary act like guessing.) There's no in-general problem with the idea of there being things that one is in no position to assert either to obtain or not to obtain (if there were, then all cases of ignorance would constitute very simple refutations of the knowledge norm of assertion). So there must be some special reason the implication is problematic in this instance. But it is not at all clear what special reason Buford thinks obtains in this case. Buford constructs (p. 117) an argument with the premise that 'what it is conversationally appropriate to assert is determined by the facts at one's disposal', along with the assumption that the 'facts at one's disposal' do not vary between low-standards and high-standards contexts. But it is not at all clear to me why a contextualist defender of the knowledge norm of assertion is committed to this combination of claims.

The second respect in which this argument is problematic is that, in taking (2) to motivate contextualism, DeRose is conflating *conversational contexts* with *practical situations*. In arguing that assertability varies with context, DeRose writes:

> It is difficult to deny that the matter of how well positioned one must be with respect to a matter to be able to assert it varies with context: What one can flat-out assert in some "easy" contexts can be put forward only in a hedged manner ("I think . . . ," "I believe . . . ," "Probably . . . ," etc.) when more stringent standards hold sway. DeRose (2002, p. 187)

The respect in which whether Ruth may assert that Fred is loyal varies according to context concerns the situation in which Ruth finds herself. If, for instance, the question of Fred's loyalty is of tremendous practical import (they're considering whether to trust him in at a crucial point in their plans), Ruth might need to be in quite a strong epistemic position to make such an assertion—in a less momentous situation (they're just trying to decide whether to throw him a birthday party), a weaker standard might suffice for assertability. But these are, in Kaplan's terms, a matter of *index*, not *context*.[28] This is merely the observation that what is assertable for Ruth, even with respect to purely epistemic norms, depends on Ruth's practical situation; it simply has nothing to do with contextualism. (It is the analogue of 'interest-relativity', in the sense of 'interest-relative invariantism', for knowledge—see §1.9.) Blackson (2004) pointed out this difficulty for DeRose's argument. In fact, the dynamic DeRose is highlighting more directly supports a kind of pragmatic encroachment, as opposed to contextualism; the argument is very similar to Fantl and McGrath's argument for encroachment, for example.[29]

The shiftiness in assertability that DeRose cites is entirely consistent with invariantism about 'knows': whether a subject knows, and whether he's entitled to assert, do vary together—but they vary with the subject's situation; this is no reason to think anything semantic varies with the attributor's context. What DeRose would need, in order to run the even a prima facie compelling analogue of his argument, would be for it to be clear that the relation expressed by 'S is entitled to assert that *p*' varies with the context of utterance. In other words, it would need to be possible for one person to express a truth with 'S is entitled to assert that *p*', and for another person, speaking of the same S and the same *p* at the same time, to express a truth with 'S is not entitled to assert that *p*'. This is not obviously the case. (Indeed, John Hawthorne seems to think this is obviously *not* the case, and runs an *objection* to contextualism along these lines.[30] (I'll examine Hawthorne's argument in §6.6 below.) And even if this *were* the case, this would only vindicate contextualism if one endorsed a particular metalinguistic generalisation of the knowledge norm of assertion—something there might well be good reason not to do.

6.5 DeRose 2009

In his (2009) update on the issues discussed in the previous section, DeRose recognises and attempts to correct for the deficiency in his 2002 argument. With much greater sensitivity to the use–mention distinction, DeRose formulates this version of the knowledge norm of assertion:

> The Relativized Knowledge Account of Assertion (KAA-R): A speaker, S, is well-enough positioned with respect to p to be able to properly assert that p if and only if S

[28] See §1.2.

[29] Fantl and McGrath (2009, Ch. 3) argue that from a normative principle linking knowledge to action (not quite the same as my KR—see §5.5), fallibilism, and the plausible idea that the propriety of actions depends on non-epistemic factors, 'impurism' follows.

[30] Hawthorne (2004, pp. 86–8).

> knows that p *according to the standards for knowledge that are in place as S makes her assertion*. (DeRose, 2009, p. 99)

I take DeRose's KAA-R to be intended to be equivalent to this formulation, which may be preferable for avoiding reification of 'standards for knowledge':

> S is epistemically entitled to assert that p iff 'S knows that p' is true in S's context.

This version of the discussion makes clear just what changes according to conversational contexts, and what changes according to the speaker's situation (i.e., the index). But, for very much the same reasons as those discussed in the previous section, this does not amount to any kind of argument for contextualism. An invariantist, who holds that what a speaker knows can vary according to her situation, can easily explain the variability of assertability without invoking contextualism. The DeRose of 2009 recognises this fact, but nevertheless suggests that there is at least a *partial* argument for contextualism here. His suggestion is that an invariantist who wants to explain the variability of assertability in this way is committed to the index-variability of standards for knowledge—only someone who holds the view he calls 'subject-sensitive invariantism' (SSI), DeRose says, can explain variable assertability without contextualism. Since he also thinks he has independent arguments against SSI, DeRose thinks the variability of assertability can play a role in an argument for contextualism along these lines: the variability of assertability, combined with KAA-R, implies that *either* contextualism or SSI is true; there are good independent arguments against SSI; so the variability combined with KAA-R implies that contextualism is true.[31]

As I mentioned in §1.9, I do not find the category of 'SSI' a helpful one. *One* reason this is so is a relatively trivial complaint about its name—*everyone* thinks knowledge depends on the situation of the subject (see Stanley (2005, p. 122))—but the deeper reason is that I don't think it picks out anything like a clear or natural set of views about knowledge. As DeRose characterises SSI, it's the view that the epistemic standards required for knowledge vary according to the subject's position. But, as I have argued, if the standards are anything like a set of relevant alternatives that need ruling out, anyone who thinks the external environment can make a difference to knowledge—anyone, for example, who thinks it's harder to know that the thing in front of one is a barn when one is in fake barn country—should accept this.[32] I think that it is a tempting mistake to suppose that contextualism and so-called SSI views have much in common; I am suspicious of the idea that they're alike in that they agree that 'different standards' apply for knowledge (. . . at different times? . . . on different occasions of use?) Here again is DeRose's statement of the truism that assertability is variable: 'What one can flat-out assert in some "easy" contexts can be put forward only in a hedged manner ("I think . . . ," "I believe . . . ," "Probably . . . ," etc.) when more stringent standards hold sway.' (DeRose, 2009, p. 187) The invariantist treatment in question has it that it's sometimes, in some situations, it's relatively easy to know, and at other times and situations it's relatively difficult.[33]

---

[31] See DeRose (2009, pp. 108–9) for remarks to this effect.

[32] My own preferred notion of standards, laid out in §1.8, identifies them with functions from indexes to sets of possibilities that determine relevant alternatives; so it is built into the definition that the standards don't change according to the index. DeRose's 'modal distance' approach might have a better chance here, but it faces its own challenges, like those pressed in Blome-Tillmann (2008) and Schaffer (2005, pp. 124–5). For related critical discussion of the degree to which SSI-style views are revolutionary, see Grimm (2011).

[33] For very similar reasons, I do not agree with DeRose that the views he calls 'SSI' are objectionable as such; I don't think they're anything *as such*. I am not sure I feel the difference DeRose hears between 'Henry doesn't know that's a barn, but if there were fewer fake barns around then he would know that's a barn' (which he thinks is fine— DeRose (2009, p. 197)) and ('she does know, but she wouldn't have known if more had been at stake' (which he thinks is unacceptable— DeRose (2009, p. 194)). But even if I did, that would be at best an argument that *stakes in particular* aren't relevant for knowledge (or that whether there are fake barns around aren't). See §1.9. For what it's worth, I'm also convinced by Jennifer Nagel that a great many of these 'now you know it, now you don't' sentences used to challenge SSI-type views follow from very orthodox

So I don't think that DeRose's 2009 attempt to resuscitate his 2002 argument for contextualism succeeds. DeRose has articulated what may, for all I've said, be a perfectly good way of understanding the knowledge norm of assertion within a contextualist framework, but I don't think there's an *argument* from the former to the latter.

6.6 Contextualism and Norms, Again

John Hawthorne argues that a contextualist treatment along DeRose's lines fails adequately to respect the motivations behind the knowledge norm of assertion. In fact, we've seen a version of his argument already, in the discussion of practical reasoning in Ch. 5. (Indeed, §2.4 of Hawthorne (2004) is called 'Assertion and Practical Reasoning' and treats the issue in these domains simultaneously.) As in the case of practical reasoning, Hawthorne complains about a disconnect between the normative facts, that are centred around the subject of evaluation, and the truths of knowledge ascriptions, which vary according to the speaker's conversational context. As Hawthorne puts it, 'owing to the purported ascriber-dependence of 'know', contextualism seems to disconnect facts about knowledge from these normative facts. And this is because the relevant normative facts do *not* seem to be context-dependent.' (Hawthorne, 2004, p. 86) Suppose that Ruth is in a skeptical context, relative to which typical knowledge ascriptions about contingent future events are false. So (1) expresses a false proposition relative to Ruth's context:

(1) Fred knows that he'll be with Mabel tomorrow.

Ruth, we may suppose, is taking seriously the possibility that some natural disaster or sudden illness will disrupt Fred's plans of being with Mabel tomorrow. But *Fred*, who is in an altogether different, and much less skeptical, conversational context, is ignoring such possibilities. He tells Mabel:

(2) I will be with you tomorrow.

Fred satisfies 'knows' with respect to (2) relative to his own non-skeptical context (i.e., (1) is true relative to Fred's context).

Hawthorne's complaint concerns Ruth's perspective on this combination of facts, combined with the knowledge norm of assertion. Because Fred is asserting only what he 'knows' relative to his own context, it seems that a proponent of KA should suppose that Fred's assertion is epistemically warranted. (Certainly this follows from DeRose's preferred (2009) version of the knowledge norm, KAA-R.) But it *also* seems to DeRose that, since Ruth could truly say (3):

(3) Fred doesn't know that he'll be with Mabel tomorrow,

she should also, given the knowledge norm, truly say (4):

---

epistemic commitments and well-established facts of psychology—see e.g. Nagel (2010). (Interestingly, Nagel herself does not take this observation to support SSI, but instead argues that the apparent implications of SSI—both its welcome ones and its unwelcome ones—can be captured and explained in a more traditional epistemic framework. Whether she's right about this further underscores one of the points I've been making—it is much more difficult than it's often realised to articulate what exactly the SSI view is supposed to be. (Nagel also says that her framework can capture, in an invariantist framework, the data typically cited in favour of contextualism, but I think it's difficult to generalise her approach to third-personal cases. See §7.1, fn. 14)

(4) Fred shouldn't assert that he'll be with Mabel tomorrow.[34]

But this, Hawthorne says, is an unacceptable combination of views. We can't hold that Fred's assertion is fine, and *also* hold that 'Fred's assertion is unwarranted' is true relative to Ruth's context. (We could do so if we embraced a kind of contextualism about 'warranted', but I am following Hawthorne in setting that option aside; see §5.3.)[35] So, Hawthorne argues, we *shouldn't* admit that (4) is true in Ruth's context, which means that (5) *is* true in it:

(5) People often flat-out assert things that they do not know to be true but are not thereby subject to criticism.[36]

This, Hawthorne says, is an unacceptable result. But I am not sure why exactly this result is meant to be unobjectionable.[37] Here is one possibility—perhaps Hawthorne is supposing that the knowledge norm of assertion is so obvious or platitudinous that any implication that sounds enough like a denial of it is unacceptable. If it's a Moorean fact that people never appropriately assert things they don't know, then the implication that (5) can express a truth might be problematic.[38] But the knowledge norm of assertion is not platitudinous in this way. I think it is an attractive theoretical posit that does a good job of explaining a wide variety of data—and even this much is controversial. It's certainly not so obviously true that one should feel embarrassed if one ends up with commitments that are only arguably consistent with it.[39]

Hawthorne also presses a different, related worry:

> Notice that our contextualist grants that there are determinate facts about when an assertion is proper. But she then goes on to divorce the question of whether a subject has knowledge from these normative facts. Isn't any such conceptual divorce strained? There is an interconnected cluster of worries here: Why should it be that 'know' is semantically highly sensitive to what is salient to the ascriber and to the interests of the ascriber, but that various related normative notions like 'what is assertable' or 'what one ought to believe'—are not context-sensitive in that way? Why should 'know' be semantically sensitive to the ascriber's interests and attention when 'what is assertable' or 'what one ought to believe' are not? Hawthorne (2004, p. 89)

[34] See the discussion on Hawthorne (2004, pp. 86–7) for Hawthorne's version of a case like this. Note that my claims about what Ruth 'could truly say' are about what is *true* in her context, not what is *assertable*. The latter would depend on her own epistemic state vis-à-vis Fred's epistemic state. Hawthorne's version of this point comes in the observation that certain sentences would be 'true in [a speaker's] mouth (whether or not [they] know it is true)'. (Hawthorne, 2004, p. 87).

[35] Although I am inclined to agree with Hawthorne that this option is not promising, I don't think the *argument* he offers to that effect is at all compelling. He writes: 'Assertability conditions and propriety conditions for practical reasoning just don't seem to vary in that way. The practical reasoning considered above is inappropriate, regardless of what an ascriber is attending to, and parallel remarks apply to the propriety of flat-out assertions of lottery propositions (in the setting envisaged).' Hawthorne (2004, p. 90) This looks like a use–mention error; the contextualist line about 'propriety of assertion' (or of 'propriety for practical reasoning') does not imply that whether the reasoning is inappropriate depends on what an ascriber is attending to, any more than contextualism about 'knows' implies that what a subject knows depends on what an ascriber is attending to.

[36] Quoted from Hawthorne (2004, p. 87), using my own numbering.

[37] I seem to have had an easier time understanding the objection in 2009, when I pressed a version of it against DeRose myself, writing that 'I worry that any principle predicting the truth, and indeed propriety, of these sentences cannot be strong enough plausibly to be identified with the knowledge norm of assertion.' (Ichikawa, 2009b) I suspect I had something like the motivation conjectured in the main text in mind; if I did, I no longer feel it.

[38] It *might* be problematic. Then again, it's a Moorean fact that people know they have hands, but many contextualists (though I am not one of them—see §3.12) think 'nobody knows that they have hands' is true in some contexts.

[39] Note also that the degree to which contextualists really are committed to these consequences is questionable. See e.g. DeRose (2009, p. 250).

While there is certainly no in-principle mystery about why it should be that pairs of words that denote related properties should differ in whether they are context-sensitive or not—the 'conceptual divorce' here looks to me like a garden-variety conceptual *difference*—I can feel at least some motivation behind a concern in the vicinity. We might put it thus: what is the theoretical explanation for why knowledge ascriptions, and the semantic profile that they have, vary in a systematically different way from the way that facts about assertability vary? It's one thing to relate assertability to the mental state *knowledge*—it's less clear how theoretically appealing the idea is that assertability is related to the satisfaction of the English verb 'knows' in a certain context. One suspects that, even if this is true, there must be a deeper explanation for it. When I develop my own proposal for how a contextualist might best implement the knowledge norm—or something reasonably close to it—one of my aims will be to explain, in a theoretically satisfying way, why the connection should work in the way suggested.

6.7 The Method of Cases

As is the case in the evaluation of other knowledge norms, philosophers engaging with the knowledge norm of assertion often propose apparent counterexamples: cases where warranted assertability and knowledge seem to come apart. And, as I emphasised in the discussion of alleged counterexamples to the knowledge norm of practical reasoning in Ch. 5, it is easy to underestimate the complexity of the task of engaging the knowledge norm in this way. In my discussion of practical reasoning, I emphasised that a separation between facts about what a subject knows, and what action a subject performs, provides room for philosophical manoeuvring, since there's no obvious, straightforward relationship between particular putative reasons and particular actions. One might think, therefore, that the case is much simpler in the case of the knowledge norm of assertion, since the relationship between the fact known and the action considered—i.e., the assertion—is much more intimate.[40] To an extent, this is correct. Nevertheless, there are complications that can go unappreciated, deriving from the uncontentious fact that *many* norms govern assertion. Even if, as Timothy Williamson suggests, the knowledge norm is the unique constitutive norm of the practice of assertion, there can be many other respects in which a given assertion can be good or bad. So a case of an assertion of a content that isn't known, but for which we nevertheless feel a sense of approbation, needn't be a counterexample to the knowledge norm. Likewise for an assertion of a content that *is* known, but that nevertheless feels improper. Evaluating the knowledge norm of assertion is a subtle matter.

I think that many of the alleged counterexamples to KA in the literature miss their target for just this reason. For example, Chris Hill and Joshua Schechter write against KA that

> [i]f a speaker has good reason to believe that p at the time of the assertion, then the speaker is not usually subject to rational criticism even if it turns out that p is false. That is to say, as far as we can see, there is no practice of criticizing such speakers for having spoken inappropriately, or for having exceeded the bounds of conversational entitlement. (Hill and Schechter, 2007, p. 109)

But as a number of authors have emphasised, the knowledge norm does not predict that there should be a practice of criticising all violations of it.[41] KA implies only that assertions that are

[40] Brown (2012a) argues against various reasons to suppose that the norms of assertion and practical reasoning must be closely connected.

[41] See e.g. Hawthorne (2004, pp. 23–4): '[The knowledge norm] is not to say that knowing p suffices for appropriate assertion. An assertion that p, even if one knows that p, might be out of place for any number of reasons: where it is irrelevant or because its implicatures mislead, or simply because one is in a conversation where p is up for debate.' See also Williamson (2000, p. 256).

not knowledge are defective qua assertion; it is a substantive assumption that all such defective assertions warrant criticism on the part of the asserter. Hill and Schechter recognise this fact, but suggest that distinguishing between norm violations and cases of blameworthiness would be needlessly complicated.[42] In my view, they underestimate the generality of that kind of need. Suppose one subscribed to a 'justified belief' norm of assertion. One would still need a story about blameless violations of that norm—suppose, for instance, that one mistakenly but blamelessly believes that one justifiably believes that p, and asserts that p. This would presumably be a case of an excusable norm violation. So the justified belief norm, as much as the knowledge norm or any other norm, would require an ability to recognise a distinction between norm compliance and excusable norm violation.[43] The knowledge norm does not create distinctive challenges here.

The same goes for cases in which non-knowledgable assertions receive the intuitive thumbs-up because of their pragmatic effects. Consider this case from Rachel McKinnon:

> Suppose that Jenny is teaching a grade 10 science class. She wants to explain the structure of an atom, and, more specifically, the electron configuration of different elements. Jenny is well aware that an early model of the electron structure of atoms, the Bohr model, is no longer considered accurate. Under the Bohr model, electrons travel in restricted *orbits*. ... More recently, though, the Bohr model has been replaced with the *valence* model. Under the valence model, due to incorporating principles of quantum mechanics such as Heisenberg's Uncertainty Principle, electron "orbits" are replaced with probability "clouds." ... So knowing all of this, Jenny also knows that her students aren't yet able to understand the valence model, but they are able to understand the Bohr model. Students of this age are typically not yet acquainted with concepts of quantum mechanics, and need to learn concepts such as the Bohr method as a stepping-stone. So when it's time to teach her students about the electron structure of atoms, she asserts, "Electrons behave according to the Bohr model." (McKinnon, 2015, p. 61)

McKinnon goes on to argue that Jenny's is a genuine assertion of a proposition she knows to be false, and indicates that she considers the assertion to be a warranted one, partly on the grounds that it is the best assertion she could have made, in terms of pedagogical effects. Now whether asserting falsehoods like this really *is* the pedagogically optimal strategy seems to me debatable. In favour of her claim, McKinnon cites Kalkanis et al. (2003, p. 63), emphasising that that students who are taught the Bohr model often leave with false beliefs about atoms—this is part of McKinnon's case for the idea that instances like this are genuine assertions of unknown contents. But Kalkanis et al. are in fact *criticising* conventional scientific pedagogy in part on these grounds; they advocate introducing elements from quantum theory alongside the Bohr model at a much earlier stage in students' science education. While I am not an expert in pedagogy, it is easy for me to believe that students who are explicitly taught that certain theories are approximations that are useful for certain purposes may end up learning and understanding the material in question as well or better than those who receive simplifying false assertions.[44] But even if McKinnon is right that the best strategy for teaching students about electrons is to make false assertions about them, this wouldn't bear in any

[42] Hill and Schechter (2007, p. 110). See also Stone (2007, p. 100) and McGlynn (2014, pp. 105–6) for similar moves.

[43] One needn't inflate reasonable norm violation into a kind of 'secondary propriety'. See Lackey (2007, pp. 604–6).

[44] To change the example to a similar one with which I have some firsthand experience, I have found, in teaching first-order logic to undergraduates, that it is *not* preferable to make the false assertion that the '$\supset$' symbol is the proper translation of the English 'if... then'. Although some textbooks make this assertion, I have found that students react much better when I tell them that it is an imperfect approximation, similar to English conditionals in certain inferential respects.

direct way on the knowledge norm of assertion; it would simply imply that, if the knowledge norm is correct, then it is useful, in giving children scientific educations, to violate the norm of assertion.[45]

6.8 KA and Good Enough Positions to Assert

So, as in the case of other knowledge norms, evaluating the knowledge norm is a sensitive matter; it's just not enough to think about whether one does well in particular cases. The knowledge norm of assertion, like other principles involving knowledge, must be evaluated on more holistic and theoretical grounds. Note that, contrary to the suggestion of Brown (2008b), KA doesn't imply either of these principles:

> NEC: one is in a good enough epistemic position to assert that p only if one knows that p.
>
> SUFF: one is in a good enough epistemic position to assert p if one knows that p. (Brown, 2008b, p. 90)

Against SUFF, consider that, just as KA is consistent with the possibility that one ought not to assert that p, even though one knows that p—because although one wouldn't violate the norm of assertion, one would otherwise do wrong—so too is it consistent with the possibility that the problem with the assertion derives from an epistemic shortcoming. The case here is exactly parallel to that in the discussion of practical reasoning in §§5.4–5.5. Suppose for instance, to take a contrived but simple case, that Samuel has made a special arrangement with Richard, whereby he won't assert that it rained yesterday unless he has good reason to believe that the police are coming. Samuel knows that it rained yesterday, but has no idea whether the police are coming. Consequently, under the circumstances, he is *not* in a good enough epistemic position to assert that it rained yesterday. His ignorance about the police prevents him from being permitted to make that assertion about the weather. So this is a counterexample to SUFF, but not to KA.

In this example I cited Samuel's insufficient epistemic position with respect to a *different* proposition as part of why he can't assert that it rained yesterday; a natural thought, then, is that Brown's suggestion just needs to be restricted to something like

> SUFF_2: one is in a good enough epistemic position *with respect to* p to assert p if one knows that p.

But this won't do either, as there was nothing essential to the kind of case given above that the propositions in question be independent in this kind of way. It may require a more fanciful case to illustrate the point, but suppose that Samuel has promised not to assert that it rained yesterday unless he knows that he knows that it rained yesterday. And suppose that he knows it rained yesterday, but he doesn't know that he knows it. Then he shouldn't assert it, and the reason he shouldn't assert it is that he lacks a particular strength of epistemic position with respect to it—he doesn't know that he knows it. But this too is consistent with KA, which would have it that an assertion that it rained

[45] Perhaps the same diagnosis should be applied to Jennifer Lackey's case of the 'creationist schoolteacher', someone who asserts the scientific consensus about human evolution even though she doesn't believe it (and so doesn't know it). Lackey (2008, p. 48) Perhaps this is a case in which it is best to violate the knowledge norm. However, there are strategies available for denying that this *is* an instance of a knowledge norm violation. I have some sympathy, for example, for the view that the schoolteacher in question is not making an assertion as an individual, but rather that she is the mouthpiece via which a broader entity (a school district, perhaps) asserts something that it knows. Turri (2014b, pp. 195–6). See also the approaches of Burge (2013), and Faulkner (2006). For discussion, see McGlynn (2014, p. 116), Carter and Nickel (2014), Lackey (2008, pp. 50–1), Lackey (2014a). The issues around group knowledge and group assertion are enormously complicated: see Pettit (2003) and the essays in Lackey (2014b).

yesterday would not violate a constitutive norm of assertion, but it would violate a promise and so shouldn't be done.

Similar cases could be made against Brown's NEC as an implication of the knowledge norm. Suppose for instance that great rewards will befall anyone who asserts something she justifiably believes, and that Edith justifiably believes, but does not know, that the servants are bringing the luncheon. Her justified belief, under the circumstances, makes it altogether proper for her to violate the knowledge norm by asserting that the servants are bringing the luncheon. So she is in a good enough epistemic position to assert that *p*, in the sense that, due to the strength of her epistemic position, asserting that *p* is the thing to do, even though this would amount to a violation of the knowledge norm.

6.9 Turri's 'Simple Test'

In the previous sections, I have argued that it is more difficult than some have thought to argue against KA via the reflection on the propriety of assertion in particular cases. But, as in the case of practical reasoning, the same goes for *defences* of the knowledge norm. Just as one shouldn't reject the knowledge norm on the basis of assertions that seem permissible even though their contents are unknown, neither can one give much *support* to the knowledge norm by observing that all-things-considered normative intuitions line up with the permissibility predicted by KA. In his (2015), John Turri describes what he takes to be empirical support for KA. Indeed, Turri describes his procedure as a 'direct test' of the knowledge norm.[46]

Turri's experiments give subjects a simple prompt, and ask for judgments about whether subjects should make particular assertions. The variable he manipulated was whether it was stipulated that the subjects did, or did not, know the contents in question. Here is one of his prompt pairs:

> Mallory manages an independent coffee shop. One of her customers is interested in the history and culture of coffee. The customer asks Mallory whether the coffee is from Columbia. Mallory knows that/doesn't know whether the coffee is from Columbia. Turri (2015, p. 391)

Subjects were asked to indicate the degree to which they agreed with the claim that 'Mallory should say that the coffee is from Columbia'. Unsurprisingly, subjects who were told that Mallory knows that the coffee is from Columbia were overwhelmingly more likely to indicate that she should say so than were those who were told that she doesn't know whether the coffee is from Columbia. Turri describes these and similar results as 'overwhelmingly favourable to the knowledge account' (Turri, 2015, p. 391).

I think it's a bit of a stretch for Turri to claim significant support for the knowledge norm, for two reasons. First, as I've been emphasizing, the question of whether the norm of assertion permits something is a different question from the question of whether something 'should', all things considered, be done. (If the vignette had stipulated that Mallory knew that the customer smelled terrible, but subjects tended not to agree with the claim that Mallory ought to say that the customer tells terrible, this would not have made for significant pressure against the knowledge norm.) The knowledge norm has it that there is a respect distinctive to assertion in which the assertion of unknown contents is defective. Turri's simple test does not get at this question; no simple test could get at this question.

Second, Turri's data is easily explained by many competitor views to KA. His data involves subjects considering thought experiments; but subjects engaging with thought experiments inevitably

[46] (Turri, 2015, p. 391)

imagine details of the scenarios that go beyond what is stipulated.[47] When subjects are told that Mallory doesn't know whether the coffee is from Columbia, typical subjects will imagine Mallory as agnostic about whether the coffee is from Columbia. (If someone doesn't know whether *p*, you'll expect them to say 'I don't know' when asked whether *p*.) One *won't* typically imagine, when told that Mallory doesn't know whether the coffee is from Columbia, that she's justifiably confident that it is, but that her belief doesn't amount to knowledge because her importer sometimes lies about origins. Consequently, when told that Mallory 'doesn't know' that the coffee is from Columbia, subjects are very likely to imagine a case in which *many* candidate norms of assertion are not met. For example, Rachel McKinnon's 'supportive reasons norm' has it that warranted assertion must be based on supportive reasons; if one is told that Mallory doesn't know, one will typically imagine a case in which supportive reasons are lacking. The same goes for many other candidate norms.[48]

Something like Turri's experiments could have gone some distance toward addressing this latter worry if he'd discovered a *stronger* correlation between stipulated knowledge and agreement with the claim that she 'should' assert than there is between the latter and other epistemic categories. But nothing like this is considered. And although I favour the knowledge norm, I am not optimistic that such tests would provide support for it. For example, I am reasonably confident that one would get similar results to Turri's if instead of stipulating knowledge, cases included the stipulation that 'Mallory is confident that the coffee comes from Columbia' or that 'Mallory is not confident that the coffee comes from Columbia'. (Or, for that matter, that 'Mallory knows that she knows that the coffee comes from Columbia' vs. 'Mallory doesn't know whether she knows that the coffee comes from Columbia'.) Without comparative results, Turri demonstrates only that there is a folk connection between appropriate assertion and the asserter's epistemic condition. This is something for which there is already wide consensus.

Although as I have been emphasising, it is a difficult and sensitive matter to decide how to interpret intuitions about the propriety of assertion in particular cases, in the next section I would like to put forward another kind of apparent counterexample to KA. Although the strategies I have been articulating so far might be extended to these kinds of cases as well, I will suggest that it may be more difficult to do so. My own suggestion will be to accommodate cases of this type by modifying our understanding of the knowledge norm.

6.10 Incremental Assertion

In his discussion of Elizabeth Fricker's (2012) treatment of assertion, John Hawthorne makes the following observation:

> [M]any cases of telling P occur in the midst of a conversation where the assertion will in part encode incremental information vis-à-vis what has already been introduced in the conversation, but may also in part encode information that has already been introduced. ... [T]here are two salient ways that her speech may turn out false: (i) the incremental aspect of the message is false, or (ii) the non-incremental aspect is false. Hawthorne (2012, p. 98)

Hawthorne goes on to observe that our evaluative response depends on which kind of false assertion one makes. Here is an example of the kind of case Hawthorne is discussing. Suppose that

[47] See Ichikawa and Jarvis (2009) for discussion.

[48] Turri anticipates a version of this worry—see Turri (2015, pp. 389–90). But he only considers the specific candidate norms that the subject believe that *p*, and that the subject 'have evidence' that *p*. Neither of these weaker norms is a major rival to the knowledge norm. Turri also does not consider norms *stronger* than the knowledge norm, like a certainty norm or a knowledge of knowledge norm; any such norm would also plausibly predict Turri's data.

Monday of this week, the pirates were prepared to attacked a ship, but decided to set it free when it turned out the ship was manned entirely by orphans. Tuesday, the same thing happened. Today, on Wednesday, they are preparing to attack a third ship, and Frederic, who has been inspecting the target's crew, says to Samuel: 'the sailors on that ship are orphans.' Samuel replies in frustration: 'the last three ships we've taken have all been manned by orphans!' Focus on Samuel's assertion here. Let us consider three possible versions of the case. In case 1, everything is as it seems, and Samuel's assertion is true and known relative to whatever the epistemic standard in question is. In case 2, the pirates are regularly duped by the sailors—the word has gotten around that the soft-hearted pirates won't attack orphans, so all three ships lied to the credulous pirates and sailed off scot-free. So in case 2, Samuel's assertion—'the last three ships we've targeted have all been manned by orphans'—is false, and so unknown. If we assume that the sailors' protestations really were sufficiently good evidence such that it was reasonable for the pirates to believe them, this will be the kind of case a proponent of KA will typically describe as a 'blameless violation' of the knowledge norm. Samuel is doing the best he can under the circumstances, by asserting something he has every reason to think is knowledge. But since he's the victim of a skeptical scenario (he's been lied to), he nevertheless violates the knowledge norm, albeit through no fault of his own.

Case 3 is the interesting one for our present purposes. In case 3 we suppose that the first two ships were indeed orphan ships, and that Samuel knows it. But Frederic is lying about today's ship. He investigated the ship, and found that many of the sailors are not orphans. (What's more, they never have been!) But, in an attempt to save the ship, Frederic lied about this fact to Samuel. Here too, Samuel's assertion is false; not all of the last three ships were manned entirely by orphans. But the *incremental* component to Samuel's assertion does everything right. Taking Frederic's assertion on board, Samuel, the cooperative conversational participant, combines Frederic's apparent information with his own knowledge, and produces something that really does follow. Although he does assert something unknown—indeed, something false—he has used his own knowledge to update the conversation in a proper way.

Hawthorne himself does not take cases like these to provide counterexamples to KA.[49] But I can certainly imagine one pressing cases like this one as counterexamples to KA. There seems to be a significant propriety to updating the informational state of a conversation with one's own knowledge, even when, in so doing, one asserts something that isn't known. To be sure, as I emphasised in §6.7, any argument from cases of proper or improper assertion against KA will be far from definitive; I certainly don't mean to suggest that the fact that we feel good about Samuel's assertion in my case 3 shows that the knowledge norm has problematic implications. But I do think there are ways in which this kind of case is more resistant than others to the kinds of strategies a defender of KA could use to explain proper norm violations. In particular, the way in which Samuel's utterance is proper seems to be a distinctively epistemic one. I don't merely mean it's proper because of Samuel's epistemic situation, as in the kinds of cases I mentioned against Jessica Brown's overly broad interpretation of the knowledge norm. Rather, the *way in which* his assertion is proper seems to be an epistemic way. Indeed, there is a sense in which Samuel's contribution to the conversation *is* an expression of knowledge.

One possible strategy to accommodate this fact, given KA, would to to suggest that Samuel hasn't *really* asserted that the last three ships were full of orphans—instead, he's asserting something more like the increment in question. For example, one might try to interpret Samuel's utterance as an assertion of the fact that *given* what Frederic said, the last three ships were full of orphans.

[49] He writes, of his version of a case like my case 3: It is very hard to see that John has a right to complain when it becomes manifest that Lizzie's speech is false, since John is fully aware that the fault in Lizzie's speech owes to faulty information that John himself has provided. I don't think any of this makes obvious trouble for the idea that knowledge is the norm of assertion. But it does suggest that the link between that norm and a right to complain may be rather complex. Hawthorne (2012, p. 99)

The strategy I have in mind has a point of connection with the 'elliptical' interpretation of assertions discussed in McKinnon (2015, pp. 56–7).[50] The idea is that the kinds of cases at issue in which the assertion might be false are implicitly offered as caveats. Adapted to my example, the elliptical strategy would have it that Samuel's assertion *really* has the content, 'the last three ships have been orphan ships, *if what Frederic just said is right*.'[51] The problem with any flat-footed version of this strategy, however, is that it's very difficult to see how an utterance of the words 'the last three ships have been orphan ships' can result in the expression of an assertion with such a conditional content. Samuel's words did not in any way suggest that he was saying anything about Fred's assertion, and so the suggestion that reference to Fred's assertion is literally a part of the content of Samuel's assertion seems unmotivated. Moreover, the elliptical strategy suggests that, in case 3, when Samuel says 'the last three ships were orphan ships', he is asserting something *true*, even though the last ship is not an orphan ship. This seems like an incredible result; certainly we would need a convincing theoretical argument to take it seriously.

Another kind of strategy in a similar spirit would be to weaken our understanding of the knowledge norm itself, holding that warranted assertions needn't themselves be known, but insisting that they satisfy some weaker, related property. For example, one might suppose that what matters for warranted assertion of p is that the speaker knows that *if the previous assertions in the conversation are true*, then p is true. This proposal would straightforwardly accommodate the propriety of Samuel's utterance above—although he doesn't know that the last three ships were full of orphans, since one wasn't, he really *does* know that the first two were. And so he knows that, if today's ship was also an orphan ship, then all three ships were. So I do think that this kind of modification has something to commend it—the proposal I will make in §6.11 will have more than a little in common with it. As it stands, however, it is insufficiently motivated. Without a theoretical explanatory rationale, the move from the knowledge norm to the knowledge-of-a-certain-kind-of-conditional norm is a bit ad hoc.

What this strategy helpfully points to, however, is the same observation implicit in DeRose's subject-centred interpretation of a contextualist knowledge norm: the details of the conversational context of the speaker play important roles in what the speaker may assert. To delve deeper into the relationship between knowledge and assertion, it will help to examine some of the theoretical frameworks available for examining conversational contexts.

6.11 Contexts and Possibilities

There are various ways to model conversational contexts. On the Kaplanian framework, a context is represented as an ordered n-tuple, providing the information necessary for the interpretation of

[50] McKinnon considers and criticises such a response to a different kind of apparent counterexample to KA, where a speaker is just wrong about the facts—it is more similar to my case 2 above. I'm much more comfortable simply calling such cases excusable norm violations.

[51] McKinnon rejects the elliptical strategy. I think she's right to do so, but I am not convinced by the Gricean explanation she offers; I don't think that Gricean implicature has much to do with the problem with the elliptical strategy; Gricean implicatures concern how speakers communicate more than what is said. But the more complex contents containing caveats, in the cases at issue, are *weaker* than the content of the declarative sentence made; so there is a straightforward sense in which they *don't* communicate more than what is explicitly asserted. In general, p entails p *unless Fred is wrong*; it doesn't *implicate* it. McKinnon doesn't explain how she thinks Grice's maxims would generate an implicature along these lines. I am also unconvinced by McKinnon's analogy to promising. She writes that '[w]hen I promise a child, for example, that I will take her to the zoo on Thursday, we understand … that I make this promise with a large number of unspoken caveats: I'll take you to the zoo, unless it's closed for repairs, or my car breaks down, or… or…' McKinnon (2015, p. 56). If she really has *promised to take the child to the zoo on Thursday*, then she breaks her promise if she fails to take her to the zoo on Thursday because her car breaks down. Perhaps she is blameless in this violation, but she didn't do what she promised to. If the promise really had included that caveat, then she would have fulfilled her promise by having her car break down.

context-sensitive language: the speaker, the addressee, the object demonstrated, the time, etc. If contextualism is correct, then the context's epistemic standard should also be included. This framework is very useful for what it's designed for: explicating the content of context-sensitive terms.

The other main approach to modelling conversational contexts is that associated with Robert Stalnaker.[52] Stalnaker's contexts are intended to track the information and presuppositions within a given conversational context.[53] Stalnakerian contexts are identified as sets of possibilities; the idea is that the possibilities in the set are those compatible with the presuppositions of the context—as Stalnaker puts it, the context set is 'the set of possible worlds recognized by the speaker to be "live options" relevant to the conversation'.[54] One might think of a Stalnakerian context as something like the *informational state* of the conversation; it represents what the speakers are taking for granted. When one makes an assertion and it is taken up in the context, the context is updated by intersecting the previous context with the set of possibilities consistent with the content of the assertion.[55]

So for example, in a context in which it is generally accepted that the men in a distant ship are pirates, the Stalnakerian context will be a set of possibilities that does not include any possibilities in which they're not pirates. A possibility in which there is no ship out there, or in which the ship is manned by marines (who are not also pirates)—or, for that matter, one in which it is manned by pirate impersonators—is not a live conversational option. But the context includes various more specific possibilities: there are possibilities where there are friendly pirates, and possibilities where there are cruel pirates. The possibilities are individuated in ways having nothing to do with the pirates, too—there are possibilities where there are pirates and July will be warm, and possibilities where there are pirates and July will be cold. The effect of an assertion, if taken up in the context, will be to *eliminate* from the context possibilities inconsistent with the content asserted. If Isabel asserts that July will be warm, and her assertion is accepted, then the context will be updated by constricting to those possibilities in previous context in which July is warm.

Stalnakerian contexts seem closely connected to the respect in which assertability is variable; they are a promising place to look for the kinds of theoretical explanations we need. Given the Lewisian form of contextualism I have embraced, knowledge ascriptions, given a context and an index, depend for their truth on a subject's relation to a particular set of possibilities—those that define the relevant alternatives for the proposition in question. If knowledge requires a particular relationship to a set of possibilities, and the respects in which conversational contexts vary is also helpfully encoded via sets of possibilities, it is natural to attempt to explain the connection between knowledge and assertability via that common theoretical framework.

6.12 The Incremental Knowledge Norm of Assertion

Consider this *Incremental Knowledge Norm of Assertion* as an implementation of a context-relative Lewisian contextualist norm of assertion:

[52] See e.g. Stalnaker (1974), Stalnaker (1998), Stalnaker (2002).

[53] Michael Blome-Tillmann's version of Lewisian contextualism, given in Blome-Tillmann (2009) and Blome-Tillmann (2014), affords Stalnakerian contexts a central role in developing an alternative to Lewis's Rule of Attention. In part for the reasons given in Ichikawa (2015), I am unconvinced by this approach, although certainly Blome-Tillmann and I have on the whole very similar approaches to contextualism. Jonathan Schaffer (2008) has also invoked Stalnakerian contexts in a way quite similar to that given in this section—my own view is quite close to his; I'll discuss his further in 6.13.

[54] Stalnaker (1998, pp. 84–5). Note that I prefer to speak in what may be a more neutral way of 'possibilities' rather than Stalnaker's own language of 'possible worlds', for reasons very similar to those mentioned in my discussion of Lewis in §1.7; I do not assume that the relevant 'possibilities' need correspond to any genuine possible world; I don't think, for instance, that all conversational contexts presuppose that Hesperus is Phosphorus.

[55] This is a simplification of Stalnaker's framework in several ways; the more complete story requires a kind of recognition or common knowledge that the uptake has occurred. These details should not undermine the suggestion of these sections, as none of the cases turn on failure to recognise or accept what has been asserted.

> IKA: An assertion that *p*, made in Stalakerian context C, is warranted only if the subject stands in the knowledge-relation to *p* relative to the set of possibilities C.

Let me explain what I mean by 'the knowledge-relation relative to a set of possibilities'. My Lewisian form of contextualism is a kind of relevant alternatives theory—to satisfy 'knows *p*' relative to a context requires a belief that *p* properly based on evidence that eliminates all relevant alternatives to *p*, where which possibilities are relevant depends on both the conversational context and on the index. (As I argued in §1.8, it is important for the index—not just the context—to play a role here, partly in order to preserve the factivity of knowledge. This will be relevant again shortly.) In order to 'know' a proposition, one's evidence must be inconsistent with all possibilities in the relevant set of worlds in which the proposition is false. The knowledge relation relevant to a set of possibilities is the relation of being such that, were that set of possibilities the relevant possibilities for the purpose of a given knowledge ascription, the subject would 'know'. So for example, suppose that Edith has as evidence that the ship is flying a pirate flag, but no evidence that would distinguish between genuine pirates and pirate impersonators. Then she stands in the knowledge relation to the proposition that they are pirates *relative to* sets of possibilities that do not include any possibilities in which non-pirates fly a pirate flag.

The reader may well wonder why a relatively simple idea need be expressed in such a complex way. Wouldn't it be easier, the reader asks, simply to talk about whether one has knowledge relative to epistemic standards? Edith satisfies 'knows' relative to contexts in which all possibilities where non-pirates fly pirate flags are irrelevant. The problem with this admirably simple approach is that it fails to appreciate the fact mentioned above—that contexts themselves do not settle the relevant sets of possibilities; indexes have a role to play as well. There is no such thing as a *context* in which all possibilities where non-pirates fly pirates flags are irrelevant. Given the factivity of 'knows' in all contexts, some such possibilities are relevant in *any* context, given an index in which one actually obtains. This is also why I am writing somewhat obliquely of 'the knowledge-relation' instead of knowledge itself. For given certain sets of possibilities—ones, for example, that include skeptical scenarios that are true—a subject can stand in 'the knowledge-relation' to *p* relative to these sets of possibilities, even though *p* is false. The knowledge-relation in question is the proper basing on (restrictedly) entailing evidence, but given the possibility of certain possible restrictions, the knowledge-relation is not factive.

But let me emphasise that I am *not* denying the factivity of *knowledge*! The kinds of sets of possibilities relative to which one might stand in the knowledge-relation to falsehoods are never those actually generated by contexts and indexes—assuming that the Rule of Actuality is an actual metasemantic constraint on which possibilities are relevant, given a particular context and index. If the ship is full of pirate impersonators, then in no conversational context will 'Isabel knows they're pirates' require only that she stand in the knowledge-relation to that proposition relative to of these nonskeptical possibilities I've been discussing.

The incremental knowledge norm of assertion is closely related to a principle that Sarah Moss has articulated:

> (EI) It is epistemically irresponsible to utter sentence S in context C if there is some proposition ϕ and possibility μ such that when the speaker utters S:
>
> (i) S expresses ϕ in C
>
> (ii) ϕ is incompatible with μ
>
> (iii) μ is a salient possibility
>
> (iv) the speaker of S cannot rule out μ. (Moss, 2012, p. 568)

Moss's idea, like mine, is that features of the conversational context play a major role in determining which possibilities one's epistemic conditions must speak to, in order appropriately to assert.[56] IKA, according to which warranted assertion requires a subject to stand in the knowledge-relation relative to the set of possibilities characterised by the conversational context (whether or not this could be equivalent to satisfying 'knows' in any context) is in important respects quite *unlike* traditional versions of the knowledge norm. In particular, IKA sometimes allows warranted assertions of falsehoods, in some cases when they are made relative to conversational contexts that are themselves inconsistent with the truth. This is just what is needed to accommodate the intuition about incremental assertion described in §6.10. There, I observed that assertions of false contents that are based on proper updating from other conversational participants' mistaken assertions, seem to be both proper and in some way knowledge-like. Fred says that today's ship is an orphan ship; Sam takes him at his word, and the context now includes only possibilities in which it is an orphan ship.

Samuel's own epistemic state rules out all possibilities in which the previous ships were not orphan ships.[57] So his evidence rules out those possibilities shaded in in Fig. 6.1. The shaded areas represent those ruled out by Samuel's evidence. There are some white possibilities outside the left circle—there are uneliminated cases in which at least one of the previous two days' shops were not orphans—but let's suppose those are rather skeptical possibilities that are properly ignored in nonskeptical contexts. Relative to such contexts, Samuel knows that the previous two ships were orphan ships.

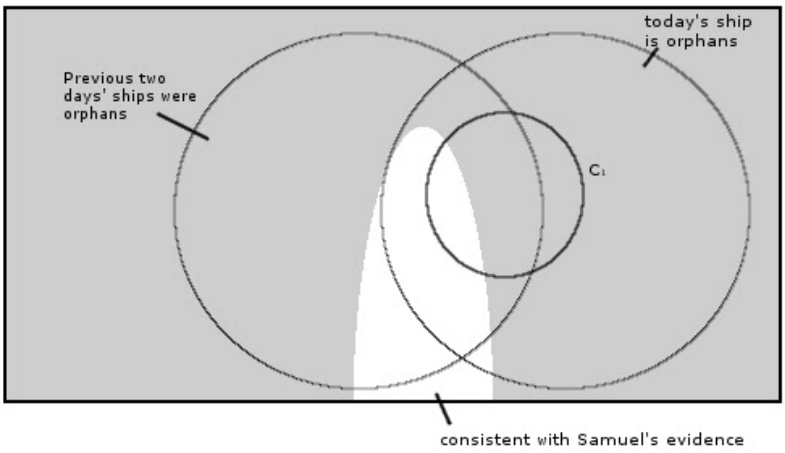

Figure 6.1: Samuel's epistemic and pragmatic situation.

After Frederic's assertion, the context is C_1. Notice that, were Samuel to assert, 'the last three

[56] A difference is that Moss's principle is in a way weaker, invoking only *salient* possibilities; mine discusses all possibilities consistent with the presuppositions of the context, whether or not they are salient. Another is that Moss does not articulate what it takes to 'rule out' a possibility. Moss points out (fn. 10) that one might be interested in deriving her (EI) from a more general source, such as a knowledge norm for assertion, but without some kind of contextualist-style shiftiness of the latter, a knowledge norm cannot predict dependence on salience. (It could predict the literal of (EI), with (iii) being an idle restriction that does no work.) Moss also points out that we want the phenomenon to be more general than assertion, extending to something like a norm of belief; the approach given in Ch. 7 is in a sense a generalisation of what I am here saying about assertion.

[57] Or at least it rules out all such possibilities that are relevant. Perhaps it doesn't rule out certain far-fetched skeptical scenarios where his investigation misled him. These are both irrelevant for the purpose of any knowledge ascriptions we're now considering, and inconsistent with the conversational context, so there's no need to worry about them.

ships were all orphan ships', he'd effectively be ruling out the right portion of C_1—that which excludes orphans the past two days. This would be the only effect on the context. What Samuel does with his assertion is eliminate from the context certain possibilities that his evidence rules out, as represented by grey in the diagrams. Samuel's operation on the context is one fully respective of his epistemic position. Fig. 6.2 shows the effect on the conversational score—it shifts the context to C_2, a proper subset of C_1, and it does so in accordance with the principle that one should remove from the context only that which one's evidence rules out.

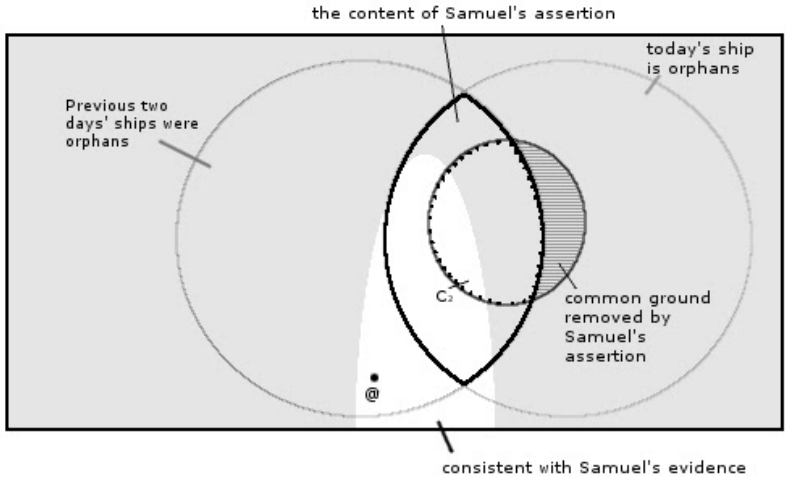

Figure 6.2: Samuel's assertion eliminates from the context possibilities ruled out by his evidence.

This, according to IKA, is what is required for an epistemically warranted assertion: that one's *incremental* assertion not go beyond one's evidence. But this is consistent with one's failure to know the truth of the proposition asserted—indeed, it's consistent with the falsity of that proposition. In the version of the case we've been considering, Frederic's assertion is false. He and Samuel are in a defective conversational context in which the actual world is presupposed not to obtain. Fig. 6.2 also reflects this defective context. So it's not true that the past three ships were all orphan ships; nevertheless, *relative to their conversational context*, Samuel stands in the knowledge-relation to that content—intersecting the context with that content eliminates an increment (the striped bit) that is inconsistent with his evidence. What IKA requires is that the striped bits contain no white.

This is the respect in which the norm is knowledge-like; it doesn't, like more familiar forms of the knowledge norm, require that the speaker know the propositional content asserted, but it does require that assertions update the conversation in knowledgable ways.[58] In asserting, one eliminates possibilities from the common ground; IKA has it that one's knowledge must rule those possibilities out.

[58] Although I won't try to draw out the connections explicitly, this approach to the knowledge norm has more than a little in common with Roger Clarke's treatment of a *sincerity* condition for assertion. On the view of Clarke (2013), an assertion is sincere if and only if the subject believes the content asserted, but *belief* is, for Clarke, characterised in a shifting, relevant-alternatives, kind of way. (I will discuss Clarke's view of belief in detail in Ch. 7.) Like my approach to knowledge and the norm of assertion, Clarke's approach to belief and sincerity emphasises and takes advantage of the structural similarity between Stalnakerian contexts and the relevant alternatives theory.

Again we can see a connection here with the conditional ellipsis proposal discussed in the previous section. What is important is that one's evidence rules out that which the assertion would remove from the context; this is very similar to the idea that one must know that *if* the context contains the truth, then the content asserted is true.[59] What Samuel must know is that if the presuppositions of the context are right, then the last three ships have been orphan ships. For the reasons articulated in the last section, I think it's a mistake to understand this as the propositional *content* Samuel asserts—but it is a good heuristic for what must, according to the norm I'm considering, be known.

It is also reasonably natural to think of the incremental knowledge norm of assertion in terms of dynamic semantics. While dynamic semantics is a broad and diverse movement, one of its central ideas is that semantic value can helpfully be modelled as *context change potential*. In the terms of the previous section, the content of an assertion like 'the last three ships were all orphan ships' can be modelled as a function from contexts to contexts, where the output context is the result of the input context after being updated with the assertion.[60] This way of thinking emphasises that *what one does* by asserting is to modify the context in a certain characteristic way. The incremental knowledge norm would put it that this modification must be a manifestation of knowledge. This would require, not that the traditional truth-conditional content asserted be known, but that the particular inferential pattern involved from moving from the initial context to the modified one be a knowledgable one.[61]

One advantage of IKA is that it allows us to capture the intuitive propriety of incrementally proper assertions of unknown contents. I don't wish to place *too* much emphasis on this fact, as we're trading in rather fuzzy intuitions that might be captured or explained away in a variety of ways, but I do think this is at least some advantage. Another, perhaps more significant, advantage is that it explains how something like the knowledge norm can fit well into a Stalnakerian picture about the dynamics of context in conversation. A third advantage of thinking of things in this way is that it can help to make sense of prima facie puzzling assertion *reports*—we seem under certain circumstances to be happy describing someone as having 'asserted that p' even though she never produced a sentence expressing the proposition that p. For example, suppose that Fred, who is on the lookout for a pirate invasion, says 'I can see several pirates making their way towards the castle'. Stanley may well report Fred's utterance thus: 'Fred says the pirates are coming now'. Under certain circumstances, this seems like a perfectly felicitous report, even though it's not at all plausible that the semantic content of the sentence 'I can see several pirates making their way towards the castle' is the proposition that *the pirates are coming now*. What is the case, however, is that relative to the presuppositions of the context, the content expressed does entail that the pirates are coming now—so his assertion had the effect of eliminating from the context possibilities in which the pirates aren't coming now (as well as possibilities in which they're coming but Fred can't see it). On the approach to the knowledge norm I am considering, that which Fred does with his assertion is to eliminate certain possibilities from the context; what he needs is for this elimination to be an expression of his knowledge.

[59] I observed in §1.13 that there is a somewhat subtle relationship between conditionals and knowledge ascriptions. I also think one can get readings like the ones discussed in that section, where the conditional acts on a restrictor on the relevant alternatives, where the relevant constraint is that 'if the context is true, S knows that p'. This could be another way of saying that S stands in the knowledge-relation to p relative to the possibility set that is the context.

[60] One of the central motivations for thinking of content in this way involves the way that assertions update contexts beyond the relatively straightforward intersection described in this chapter. For example, it is helpful for making sense of the interaction of definite and indefinite descriptions and pronouns across sentences—i.e., 'A man comes in. He sees a dog.' (van Eijck, 2010, §4) But these subsentential details won't be particularly relevant for the discussion of this book.

[61] Enoch and Schechter (2008) suggest that inferential moves, rather than propositions, are the primary bearers of central epistemic properties.

6.13 Schaffer on Contrastivism and Assertion

Jonathan Schaffer is a contextualist who has developed an approach to the knowledge norm similar to the one explored here.[62] Like me, Schaffer emphasises the role of assertions in restricting Stalnakerian contexts, and draws a connection to his own relevant-alternatives approach to contextualism. Unlike me, however, Schaffer emphasises the theoretical role of the *question under consideration*; as Schaffer puts it, the question under consideration imposes a more specific structure to the set of possibilities that are live in the context, grouping together those possibilities that amount to the same answer to the question. So for instance, in the case I've been discussing, in addition to the presuppositions of the conversational context that determine which set of possibilities is live—it's common ground that there's a ship out there, for instance, so no possibility in which there's no ship is part of the context—there is also, as part of the context, a particular question under consideration: perhaps a question concerning what kind of a ship it is. Questions, in Schaffer's framework, are partitions of the live possibilities into subsets that amount to *answers*.[63] So one question might be interested in dividing the space of possibilities into those in which it's a pirate ship, a naval ship, or a trading ship; it adds structure to the Stalnakerian context by grouping all possibilities in which it's a pirate ship together, and so on for the other answers. Another more demanding possible question under consideration might have it that 'pirate ship' doesn't counts as a full answer—perhaps one would need to identify the ship as belonging in particular to the Pirates of Penzance to answer this more demanding question. (But even to answer this more demanding question, one needn't distinguish possibilities in which the ship contains an even number of the Pirates of Penzance and those in which it contains an odd number of them.)

While I do not doubt that questions under consideration play significant pragmatic roles, and that there may well be good reason to work with a notion of context rich enough to include them, I do not see that Schaffer is right to suppose that this additional structure plays a central role in articulating epistemic norms of assertion. Schaffer's suggestion is that we can understand norms of assertion by reference to the fact that assertion aims at answering the conversational question. On Schaffer's approach to knowledge, to know p is always to know that p is the answer to some question, and some questions make doing so harder than others. (To use one of Schaffer's own examples, it is easier to know that a given bird is a goldfinch *rather than a raven* than it is to know that it's a goldfinch *rather than a canary*.) Given these two approaches, Schaffer reasonably explicates the knowledge norm of assertion by connecting them: S may assert that p in a context defined by question Q if and only if S knows that p is the answer to Q. (Schaffer, 2008, p. 10) There is certainly a similarity between Schaffer's approach and my own, but I do not prefer Schaffer's approach, for three reasons.

First, I think that requiring that warranted assertions be answers to the question of the conversation places too strong a topicality constraint to play the role of a constitutive norm of assertion. I think that some assertions that are not answers to the conversational question are nevertheless permitted by the relevant norms. Perhaps it is rude to assert irrelevantly and change the subject, but this isn't a matter for a constitutive assertion norm. (There is no contradiction in the idea of a practice of assertion in which people don't care about addressing the questions already under consideration.)[64]

[62] Schaffer's 'contrastivism' satisfies 'contextualism' as I use the term, because it implies that sentences of the form 'S knows p' express different propositions in different conversational contexts. Schaffer himself sometimes reserves 'contextualism' for the more specific linguistic thesis that 'knows' is an indexical, which he denies. (Schaffer, 2004). On Schaffer's view, 'knows' always contributes the same semantic content, but it is a predicate with an additional argument for the question under consideration. (Schaffer also sometimes uses the weaker notion of 'contextualism' that I am using—as he does in Schaffer (2008, p. 14) where he calls contrastivism 'an unusual version of contextualism'.)

[63] Schaffer inherits the framework from Calrlson (1983), Ginzburg (1996), and Roberts (2004). See Schaffer (2008, §1.2).

[64] Schaffer's view may be too strong in another related way too—suppose that one knows that p (relative to whatever question you like), where p isn't *irrelevant* to the question, but isn't an answer to it either. Perhaps it's only a partial answer (ruling out a few of the possibilities), or perhaps it's a relevant ancillary claim ('I could get a better look if I climbed on

Second, it is unclear how to think of the status of factivity in Schaffer's framework. Given the fact that conversational contexts can include false presuppositions, sometimes the question under consideration will have no true answer. This can occur in cases in which an articulation of the interrogative itself carries a false presupposition, as in the infamous 'have you stopped beating your wife?', but also in more subtle cases in which something like the question itself is perfectly fine, but the partitioning of possible answers that constitute the context does not include the truth.[65] For instance, consider a context in which the question under consideration is, 'what made that sound I just heard?', and the possible answers are 'invaders' and 'the wind'. Was it invaders who made that sound, or was it the wind? If as a matter of fact, it was a policeman who made the sound, even though all parties to the conversation assume it wasn't, then the context does not include any true answer to the question. Schaffer's question-relative knowledge norm then will imply that no assertion in this context is permitted. Assuming that Schaffer's contrastive approach to knowledge is factive—that one can't know that p rather than q in cases where both are false—then nothing will be assertable in such a context.

Third, relatedly, Schaffer's approach does not explain what is appropriate about the kinds of incremental assertions discussed in §6.10. Suppose that the question under consideration is a partition of the Stalnakerian context, grouping together those possibilities in which exactly one, exactly two, and exactly three of the last three ships were an orphan ship. (Colloquially, this would be the question, 'is it one, two, or three of the last three ships that were orphan ships?')[66] Samuel asserts—warrantedly, I suggested—that all three of the last three ships were orphan ships. This assertion is counted as warranted by the incremental knowledge norm because the way that he modifies the context is a knowledgable one: he has updated it precisely by ruling out some possibilities that his evidence ruled out. But—again, assuming the factivity of Schaffer's contrastivist treatment of 'knows', Samuel does *not* know that *three rather than two of the last three ships were pirate ships*. So Schaffer's view, like the traditional KA, implies that these 'incrementally warranted' assertions are unwarranted.

6.14 Explaining Moore-Paradoxicality

One of the central virtues claimed for the knowledge norm of assertion is that it provides a good explanation of the unacceptability of the Moore-paradoxical 'p but I don't know that p'. Although everyone should admit that the contents expressed by such sentences are sometimes *true*, it seems like they can never be appropriate things to *say*, and the traditional form of the knowledge norm, according to which assertion that p requires knowledge that p, explains why. The proposal under consideration, IKA, is also well-able to explain this phenomenon. Although there *will* be possible cases in which one doesn't know that p but can nevertheless warrantedly assert that p in certain conversational contexts, this will only occur in cases in which the context is inconsistent with the actual not-p world. In such cases, one will typically not be in a position to know that one is in such a case. As in the case of Samuel and Fred, one will have accepted some falsehood, either because it's been asserted or because it was already a presupposition. If one knows that it's false that yesterday's ship was full of orphans, one won't typically allow the conversational context to be updated with this

that rock'). Schaffer recognises a role for such relevant assertions that aren't answers earlier in his paper—see e.g. Schaffer (2008, p. 3)—but when he articulates his version of the knowledge norm, he does not seem to accommodate them.

[65] I'm trading on a bit of an ambiguity on 'question' here—in the ordinary sense of the word, the question is fine—in the example that follows, there is nothing wrong with asking what made that sound. On the more technical notion Schaffer is employing, however, the question is identified with the set of possible answers; uttering 'what made that sound?' will ask a different question when different answers are options.

[66] There could be a question that treats zero an option as well, but Fred's and Samuel's presumably isn't, since Fred has just asserted that the last one was one, and the context has been updated accordingly.

proposition. So in ordinary cases in which p is assertable but unknown, the fact that it is unknown will itself be unknown. More to the point, that fact will not be entailed by the intersection of the context and the subject's evidence. And so it won't, on IKA, be assertable.

Admittedly, it is *possible* to know that one's conversational context is defective in this sort of way. In such cases, the straightforward explanation of Moore-paradoxicality won't hold. But alternate explanations are available in these cases. These will be cases in which, for whatever reason, one is knowingly *allowing* the context to exclude the truth. For example, consider a modification of the case we've been working with, where Samuel knows that Fred is trying to deceive him. As before, Fred lies and says that today's ship is an orphan ships, but now, Samuel knows that Fred isn't to be trusted on this score. But he doesn't want to led Fred know that he's on to him, so he plays along, allowing Fred's assertion—which he doesn't believe—nevertheless to update the conversational context as before.[67] This, then, will be a case in which IKA implies that (a) Samuel may assert that the last three ships were orphan ships, and (b) Samuel knows that he doesn't know that the last three ships were orphan ships. Depending on the details of the case, Samuel might even know that they *weren't* all orphan ships—or perhaps he merely knows not to take Fred's word for it, and so remains agnostic about its truth. So the standard KA explanation for the infelicity of 'p but I don't know that p' doesn't seem to apply in this kind of case. Nevertheless, I think the IKA theorist has available strategies for explaining the impropriety of such Moore-paradoxical assertions, even in cases like these.

Notice, first, that if as we've supposed, Samuel wants to play along with Frederic's assertion, then asserting the Moore-paradoxical sentence in question would be practically disastrous for him. For his assertion that he doesn't know that all three ships were orphan ships implies that *either* he's violated IKA—he doesn't know that if today's ship is an orphan ship, then all of the last three were—or that he thinks he didn't acquire knowledge of what Frederic just told him. But to communicate that one doesn't know what one has just been told, while true in this instance, would be to give away the fact that he's suspicious of Fred's testimony. And if he communicates that he doesn't believe Frederic's assertion, then *ipso facto*, he's not treating it as common ground. So on the Stalnakerian framework, the context is *not* limited to possibilities in which today's ship is an orphan ship, and so his incremental assertion will rule out possibilities that are consistent with his evidence.

So it seems that a more complicated story can explain what is wrong with Moore-paradoxical assertions, even in the rare cases in which one knows that one's context is defective. Asserting that p in such cases requires participation in the defective context, while asserting that one doesn't know that p amounts to destroying it.

[67] On Stalnaker's view, the context is equivalent to that set of propositions for which the conversational participants are disposed to treat as common ground; it is not required that individual subjects actually believe them. See Stalnaker (1974, p. 49), Stalnaker (2002, pp. 716–17).

Chapter 7

Belief

According to an attractive and popular idea, belief is by its nature *teleological*; to believe something is to be in a certain kind of normative situation. Metaphorically, beliefs have *aims*, and they are successful insofar as they achieve their aims, and unsuccessful insofar as they do not. One version of this idea that has attracted many epistemologists is the idea that belief aims at *truth*; a true belief is a belief that is good or successful *qua* belief; a false belief is one that fails in its distinctive way.[1] In recent years, a number of epistemologists have suggested, in opposition to this familiar picture, that *knowledge*, rather than truth, is the characteristic aim of belief.[2]

As in many cases discussed throughout the book, the knowledge norm of belief calls out for clarification, given a contextualist framework. One who endorses a knowledge norm for belief thinks that there is a sense in which 'a belief is an appropriate one if and only if it is knowledge' is true; but if one is a contextualist, one has many choices about the degree of metalinguistic generality with which one embraces it.[3] I've laid out the shape of the options here several times already. In a way very parallel to John Hawthorne's challenge of knowledge norms of assertion and practical reasoning, Timothy Williamson has argued that the plausibility of a knowledge norm for belief makes trouble for contextualists. Williamson offers this case:

> On the basis of memory, Mary believes truly that she had her purse yesterday morning; background conditions are normal. In an ordinary context, Speaker(low) says 'Mary knows that she had her purse yesterday morning'. Simultaneously, in a seminar on scepticism, Speaker(high) says, of the same person, 'Mary does not know that she had her purse yesterday morning'. The contextualist insists that both speakers speak truly. Suppose that both speakers endorse and apply the argument of the preceding paragraphs, as uttered in their respective contexts. Thus Speaker(low) says 'It is not wrong for Mary to believe that she had her purse yesterday'. Speaker(high) says 'It is wrong for Mary to believe that she had her purse yesterday'. Both speak truly only on a contextualist account of 'wrong'. The relevant sense of 'wrong is doubtless not

[1] See Williams (1973, p. 148), Whiting (2013), and the essays in Chan (2013).

[2] The clearest commitment to such a view I know is in Miracchi (2015a) and Miracchi (2015b); precursors can be found in Williamson (2000, pp. 47, 255–6).

[3] 'Appropriate' here does not mean 'all things considered appropriate'. There are many different norms that govern beliefs; not all knowledgable beliefs are appropriate *in every sense*. Perhaps for instance there are moral reason not to believe certain truths that one is in a position to know— Driver (2001) suggests there may be cases like this. (So pointing to beliefs that it is in some sense bad to have, but that are knowledge, is no refutation of the knowledge norm; compare my parallel discussion of assertion in Ch. 6.) I also think there are other *epistemic* norms of belief—indeed, a central point of Ch. 4 was to argue that there is an internalist norm on belief that falls short of knowledge. (However, satisfaction of these norms is entailed by satisfaction of the knowledge norm.)

> specifically moral. Nevertheless, this contextualist resolution of the apparent disagreement still seems glib and superficial. Since epistemic standards vary so wildly between the ordinary context and the epistemological context, they vary markedly between at least one of those contexts and the agent's context: Mary's. Yet the primary question seems to be 'Is it wrong for me to believe that I had my purse yesterday?' as uttered by Mary. The epistemic standards relevant to answering that question are those operative in Mary's context. (Williamson, 2005b, p. 110)

Williamson thinks that it is unacceptable to offer a conciliatory contextualist diagnosis of apparent normative disagreements like whether it is appropriate for Mary to believe that she had her purse yesterday. No doubt there is some context-sensitivity in the relevant normative language, but it is not, according to Williamson, plausible to suppose that this fact is the key to dissolving the dispute between our two speakers.

As I indicated in the Introduction to this book, I think it would be a mistake to assume from the outset that contextualists should respond to every version of this kind of charge in the same way. I responded, in Ch. 5, to the practical reasoning version of this complaint by arguing that the speakers in question needn't have even a superficial disagreement about what the subject ought to do; although contextual differences make for a difference in whether 'knows' is satisfied, they needn't make for a difference in whether 'ought to Φ' is satisfied. In Ch. 6, I suggested that the proper response was to modify the knowledge norm in a way that gave fundamental importance to the conversational context of the speaker herself; someone in a context suitably dissimilar to hers may speak truly with 'she didn't know *p*', but would be *mistaken* to infer 'she was wrong to assert *p*'. (And so likewise for someone in a less skeptical context, being mistakenly permissive.) Here, I will pursue a different strategy. Unlike in the case of assertion, I will suggest that the relevant knowledge norm really does hold in metalinguistic generality. I will allow that 'belief aims at knowledge' is true in *every* context, so I will follow Williamson in allowing that there could be speakers in different contexts, one for whom 'it is wrong for Mary to believe that she had her purse yesterday' is false, and one for whom it is true. Nevertheless, I also want to agree with Williamson that a contextualist diagnosis of the normative language at play here is not a plausible resolution of the apparent disagreement in cases like this one. As far as I can see, this leaves only one option. If the assertion that 'it is wrong for her to believe *p*' is true in the mouth of someone in a high-standards context, and an assertion of 'it is not wrong for her to believe *p*' is true in the mouth of someone in a low-standards context, one can't be a straightforward invariantist.[4] So I face pressure to be a contextualist about 'it is wrong for her to believe *p*'. The way to do this without a contextualist treatment of the normative language (i.e., they're using 'wrong' to pick out different properties) is to embrace contextualism about 'believes'.

To forecast the kind of approach I'd like to defend, if we suppose that Mary is responding correctly to her available evidence, her doxastic state will be one that satisfies 'believes' relative to the low-standards contexts where she also satisfies 'knows'; relative to more demanding contexts, 'Mary believes' will be false—and so will the corresponding knowledge ascriptions. So the high-standards speaker can express a truth by 'it is wrong for her to believe *p*', without committing to any criticism of Mary, who *doesn't* 'believe *p*', while at the same time the low-standards speaker can express a truth by 'it is not wrong for her to believe *p*', perhaps while recognising that she *does* 'believe *p*' and that her doxastic state is perfectly fine. It's not difficult to see how an approach like this one avoids Williamson's critique; I also think there is some reason to treat it as independently motivated.

The plan for this chapter is as follows. I'll begin with some methodological remarks and clarifications about the scope of the contextualist approach offered. Then I'll try to make the case that there

[4] One could offer a *relativist* treatment of this data; see e.g. MacFarlane (2005), Brogaard (2008). But I will continue to set relativism aside.

is data that is well-explained by a kind of contextualism about 'believes'. In §7.3, I'll introduce and explain one of the main inspirations for my approach, that defended by Roger Clarke; then I'll go on to spell out the contextualist interpretation of the approach to doxastic states that I like, and connect the resultant approach back to the knowledge norm of belief.

7.1 Outright Belief

I begin with a clarification. The sense in which I think it's plausible that there is a knowledge norm on belief concerns what epistemologists sometimes call *outright belief*. The ordinary language most people use to talk about doxastic states is messy, and many apparent belief ascriptions discuss something weaker than outright belief. There is a sense, for instance, in which anyone who assigns a relatively high probability to some event might be said to 'believe' that it's going to happen. When the home team is ahead by four runs going into the ninth inning, there is a sense in which I will typically 'believe' that they're going to win—on the rare occasion when they go on to lose, I find this surprising. But this kind of high credence is not outright belief in the sense relevant for the knowledge norm. (I doubt that I count as 'knowing' that the home team will win, in such situations, even relative to very low-standards contexts—but I don't think the knowledge norm implies I've done anything wrong in cases like this.) Outright belief is a fuller commitment to the truth of the content believed. If I have a high credence in some proposition, but it turns out that, although I did proportion my credence well to my evidence, this was a surprising case where the proposition turns out to be false, my high credence wasn't *mistaken*. But if I *outright believe* that p, where p is false, then even if I had very strong evidence in support of my belief, my belief was wrong.[5] Outright belief, in the relevant sense, is something much closer to an attempt at knowledge.

Not all epistemologists recognise a category of outright belief that works in this kind of way. Dustin Locke, for example, assumes that high credence suffices for belief when arguing for a normative distinction between believing and 'premising', which is roughly relying on a given proposition as established for the purpose of practical reasoning. He describes a case in which a subject has justified high non-maximal credence in some proposition, but where it would be irrational to bet on that proposition, given a certain kind of extreme payoff structure. To rely on the proposition would be an error that Locke calls a 'premising mistake'. He goes on to write: 'Are premising mistakes really distinct from doxastic mistakes? They are. Suzy does not make a doxastic mistake, at least not with respect to the proposition that Z will win. As we said, the evidence strongly supports the proposition that Z will win, and so Suzy makes no mistake in *believing* that Z will win. Her mistake is in allowing this belief to figure into her practical deliberation in the way that it does.'[6] In assuming that stipulating evidence that justifies a high credence suffices for establishing justified *belief*, Locke sets aside the category of outright belief that this chapter is about. In a similar way, Aidan McGlynn argues against knowledge norms for belief, based in part on apparent cases of 'belief' that the subject knows to be insufficiently supported to amount to knowledge; but the cases he is considering are probably not outright beliefs in the sense of this chapter.[7] For example, as I write this passage in August 2015, the evidence available to me makes it very probable that Hillary Clinton will be the Democratic Presidential nominee in 2016.[8] But I think it's very implausible that I *know* that she will be the nominee. Even if she will—and even if my evidence is actually much stronger than I think it is—I am insufficiently doxastically committed to this proposition to 'know' it, even relative to quite lax conversational contexts.

[5] Compare the 'truth standard' of Fantl and McGrath (2009, p. 141). See also Goldman and McGrath (2014, p. 262).

[6] Locke (2015, p. 75, emphasis in original)

[7] (McGlynn, 2014, pp. 28–9)

[8] As I write, Nate Silver of fivethirtyeight estimates her odds at 85%. http://fivethirtyeight.com/features/hillary-clinton-scandal-inevitable-problems/.

Sometimes when I make these clarifications, philosophers accuse me of using language in a revisionary way, stipulating an unrelated notion of belief to fit the knowledge-first program. For example, McGlynn (2014, p. 29) writes that 'the claim that we don't believe lottery propositions seems ad hoc and in conflict with the self-reports of many philosophers (including myself) and non-philosophers.'[9] While I admit that there is some truth in the thought that ordinary language belief ascriptions do not always pick out the notion of outright belief I'm discussing, I do not think that my proposal is broadly revisionary in the way McGlynn suggests. I do not think that philosophers and laypeople always speak falsely when they self-attribute 'belief' in cases where there is no outright belief. I am happy to countenance a use of the language according to which, for example, 'Jonathan thinks that Hillary Clinton will be the Democratic nominee' is a true description of my doxastic state in August 2015. Sometimes we do use 'thinks' or 'believes' to signal this relatively weak kind of attitude of finding-likely.[10] So I don't think that my proposal implies, as McGlynn suggests it would, that I need take issue with philosophers who self-ascribe 'belief' in cases like this. But I do think that this is a different sense from the notion of *outright* belief that primarily concerns me.

Second, focus on a doxastic notion of *commitment* that goes beyond high credence is not ad hoc. I take myself to be following in a significant epistemological tradition of treating belief as a kind of full commitment; contrary to McGlynn's suggestion, this is not an idiosyncratic quirk of the knowledge-first movement. Consider lottery propositions—for example, supposing that I hold a ticket in a one-in-a-million fair lottery, consider the proposition that my ticket will lose. I am justified in holding a very high credence—.999999—in that proposition. But epistemological orthodoxy has it that I do not know it, even if it is true. In any sense of 'belief' for which high credence is sufficient, orthodoxy implies, therefore, that lottery propositions are justified beliefs that do not amount to knowledge. (I take it the belief should be counted as justified, on this view, if the high credence that constitutes it is itself justified.) So orthodoxy and the high-credence-as-belief view imply that lottery cases are a kind of Gettier case, in the sense that they could have been listed alongside Gettier's own cases as refutations of the idea that justified true belief suffices for knowledge. In this context it is somewhat remarkable that lottery cases were not considered particularly relevant in the analysis of knowledge literature.[11]

Indeed, Gettier (1963) himself provides a telling clue about the notion of belief that epistemologists have been working with. On the first page of that paper, before the famous second page and its influential counterexamples, Gettier articulates the views about knowledge that he intends to target. I don't think epistemologists always remember this fact, but Gettier laid out *three* approaches to knowledge, which he took to be, for the purposes of his counterexamples, effectively equivalent. They are:

> (a) S knows that P iff (i) P is true, (ii) S believes that P, and (iii) S is justified in believing that P.
>
> (b) S knows that P iff (i) S accepts P, (ii) S has adequate evidence for P, and (iii) P is true.

[9] McGlynn also offers a separate critique of this strategy—namely, that 'the distinction between fully believing something and merely believing it to a high degree is vague.' McGlynn (2014, p. 29) One of the projects of this chapter is to explicate more specifically how the notion of outright belief might work.

[10] Indeed, in some contexts, we use this kind of language to signal *relatively* high credence that is nevertheless *quite* low—less than .5, even. When sports writers submit their predictions about who will win the championship in an upcoming season, for example, they're naming the team they think has the best chance. Nate Davis of *USA Today* has predicted that the Green Bay Packers will win the Super Bowl in the coming NFL season. http://www.usatoday.com/story/sports/nfl/2015/08/04/record-projections-super-bowl–50-packers-colts/30985171/. I assume his credence in this proposition is considerably less than .5, though it is higher than his credence that any other team will win. Nevertheless, there may well be a true reading of 'Nate Davis believes that the Packers will win the Super Bowl this year'.

[11] See McGlynn (2014, p. 25) on this point.

> (c) S knows that P iff (i) P is true, (ii) S is sure that P is true, and (iii) S has the right to be sure that P is true.

Gettier focuses on (a), with the intention that it should stand in for all the kinds of views canvassed. He sees all three formulations as ways of articulating the idea that knowledge is equivalent to justified true belief. So it is that we speak of the 'justified true belief' theory of knowledge. But only definition (a) uses the word 'believes'; the other two definitions articulate the thought in question with 'accepts' and 'is sure that'. In colloquial English, 'believes' is very different from 'is sure that', but Gettier here seems to be tapping into an epistemological tradition in which they're effectively synonymous.[12] The fact that this could be done without particular remark further attests to the somewhat idiosyncratic or theoretical sense in which epistemologists discuss 'beliefs'. It is also relevant, in this context, that folk belief ascriptions almost never use 'belief' or its cognates. The colloquial English way to ascribe beliefs is using 'thinks that' language.[13]

The epistemological notion of belief in which I'm interested is to a considerable degree a theoretical one. It is something much more like 'being sure'—it is taking the question to be in some sense *settled*. As Fantl and McGrath (2009, p. 141) put it, 'if you believe that *p*, then your mind is made up that *p*.' Jennifer Nagel recognises this fact in her discussion of 'epistemic anxiety' as a force that interferes with knowledge. As she puts it, '[s]ometimes, belief formation is automatic and effortless; at other times—and notably, in high-stakes circumstances—it can be harder to make up one's mind.'[14] Implicit in the contrast is the idea that *making up one's mind* is necessary for outright belief.[15] There is also a case to be made— Myers-Schulz and Schwitzgebel (2013) make

[12] No doubt, Gettier played some role in *creating* the tradition in question—see Kaplan (1985) and Dutant (2015).

[13] When I tell laypeople that I am working on the nature of belief and the question of which things it is reasonable to believe, they typically assume that I'm talking about religious matters.

[14] Nagel (2010, pp. 413–4). Nagel also uses this observation in an attempt to debunk patterns of shifty intuitions that are often thought to motivate contextualism or some kind of interest-relative invariantism. Although that idea isn't central in my present context, since it is at odds with many of the motivations throughout this book, I'll discuss it briefly here. Nagel's idea is that, when contemplating the costs of being wrong, human subjects have a tendency to seek out more information before settling their minds on a question. So in high-stakes situations, subjects *stop believing* many of the things they previously believed. Nagel attempts to use this phenomenon to explain variable patterns of epistemic intuitions without positing anything like a change in the epistemic standards operative; epistemic intuitions change because the epistemic *facts* change: subjects stop knowing, because they stop (outright) believing. While this strategy has considerable appeal—and Nagel backs up her psychological speculations with considerable empirical support—I do not think it is general enough to undermine the case for contextualism, as it is only applicable to first-person cases. If you and I are each discussing Constance, and I am experiencing epistemic anxiety but you are not, Nagel's strategy offers no way for your knowledge ascription, and my knowledge denial, both to be true. Instead, she offers an error theory whereby I mistakenly project my own epistemic anxiety onto Constance, *wrongly taking her* to be experiencing doubt inconsistent with outright belief. Nagel (2010, p. 413). Although she doesn't consider hers to be any kind of pragmatic encroachment or IRI/SSI theory, her view does behave like that of theorists like Stanley, Hawthorne, and Weatherson in this respect. See e.g. Nagel (2010, p. 413) and Stanley (2005, pp. 101–03) for similar manoeuvres. (The pragmatic encroachment of Fantl and McGrath (2009) does not engage in these manoeuvres; they invoke contextualism instead.) Contextualism requires no such error theory. (But contextualism is consistent with Nagel's suggestion that one reason subjects don't know when they are in high-stakes situations is that they are experiencing doubt inconsistent with belief.) The similar strategy of Gerken (2013) may have stronger prospects on this score, as it arguably depends less on first-personal cases.

[15] Compare remarks by John Gibbons:

> 'when you do make up your mind and believe that *p*, then you are taking a stand on that issue'. Gibbons (2013, p. 201)

and Timothy Williamson:

> "'Believe" here is used in the sense of outright belief; assigning a high subjective probability (short of 1) to the proposition that P does not suffice for believing that P. For example, in believing on statistical grounds that my ticket has a chance of 0.9999 of not winning the lottery, and consequently assigning a subjective probability of 0.9999 to the proposition that my ticket will not win, I do not have the outright belief that my ticket will not win, for that would involve outright belief in the crucial premise of an argument for not buying the ticket in the first place (however large the prize).' Williamson (2005b, p. 108)

it—that there is an ordinary notion of 'belief' that is *stronger* than the one epistemologists focus on: they offer cases of apparent knowledge *without* belief, as in cases where one 'knows full well' that student-athletes are no less intelligent than other students, but due to prejudice nevertheless 'believes deep down' that they're inferior.[16] I have no quarrel with their claim that certain reasonably widespread intuitions support 'knowledge' without 'belief', but I deny that this constitutes a challenge to the epistemological orthodoxy that knowledge requires belief; that orthodoxy is employing a rather theoretical notion of outright belief, about which ordinary linguistic intuitions have little to say. In the remainder of this chapter the term 'belief' will be reserved for outright belief.

So my methodological stance about belief is more similar to that about justification than to that about knowledge. As I explained in the discussion of justification in Ch. 4, I took myself there to be offering a kind of *explication* of a useful philosophical notion; it is the same kind of project I am engaging with here. I wish to offer a proposal about the best way to understand epistemologists' talk of belief in the sense in which it's used in the literature on e.g. 'justified true belief'. As such, the evaluation of my proposal is a more complicated matter than that of theories purporting to analyze ordinary notions. It is to some degree answerable to intuitions about whether 'belief' is present in particular cases; but it also matters how well it fits into our broader epistemological theorizing—including normative connections to knowledge.

7.2 Shifty Data

Once it is clarified that the notion of belief of interest is that of outright belief, some of the motivations for a contextualist or otherwise 'shifty' approach to belief are readily seen. As I've been pointing out, high credence does not suffice for belief. But with this observation in hand, it is not difficult to motivate an argument to the effect that none, or perhaps almost none, of rational people's doxastic states constitute belief. The argument runs like this: grant that high credence isn't enough for belief; one needs to be fully committed to a proposition to count as believing it. But for all or almost all doxastic states, there is inevitably room for further commitment. Consider again lottery beliefs; if Mrs. Partlett has one of the two hundred tickets for the village raffle, her credence that she will not win is quite high—if she is rational, it's .995. But since she *leaves open* in her mind the possibility that she will win, she doesn't count as *believing* that she will lose. She recognises an open possibility that she will win. But—the argument continues—insofar as she is properly epistemically humble, *nearly all* of her apparent beliefs are like this. One would ordinarily describe her as believing that there will be a banquet this evening, but if you ask her whether there's any possibility it might be cancelled at the last minute, she'll demur. Maybe even likewise for her belief that Constance is her daughter—if you ask her whether there's any possibility that she might accidentally have been switched at birth with another child, perhaps she'll say she's not sure. (Perhaps she wouldn't—many ordinary people would rule out such far-fetched possibilities. Let's suppose that Mrs. Partlett doesn't. We'll return later to the question of what to think about more dogmatic people who do.) So maybe Mrs. Partlett has almost no beliefs at all.

This argument has obvious similarities with familiar skeptical arguments against knowledge. Just as the familiar skeptic emphasises distant uneliminated not-p possibilities in order to undermine the thought that one knows p, so too does our *belief*-skeptic emphasise just the same kinds of possibilities to argue that the subject doesn't even *believe* that p. Find a not-p possibility about which a subject is agnostic, and you've shown that the subject doesn't *really* believe that p. But there is another point of similarity with the knowledge case that counts in the opposite direction: many

[16] Myers-Schulz and Schwitzgebel also looked at cases like Radford (1966)'s influential case of the 'unconfident examinee'—this is a case about which I'm more inclined to say the inclination to attribute knowledge is in error.

nonskeptical belief ascriptions, just like many nonskeptical knowledge ascriptions, are highly intuitive. For example, it seems very plausible that Mrs. Partlett *does* believe that there will be a banquet this evening—attribution of that state explains why she is telling her neighbours that there will be a banquet this evening, and why she's currently dressing her daughter for the banquet. Now admittedly, I must tread carefully here, given the clarifications emphasised in the previous section. 'Belief' as I am using it is a philosopher's term of art; I cannot look to the folk to corroborate my claim that it is intuitive that Mrs. Partlett believes this, the way I *can* look to them for corroboration of the intuitiveness of knowledge ascriptions. The data at our disposal are more limited. Nevertheless, there is good reason to regard the idea that no one (outright) believes anything as radically revisionary. Philosophers often cite beliefs in the explanation of action.[17] Philosophers almost all hold that belief is necessary for knowledge, and that many of us know many things. So I do think there is considerable pressure against the skeptical conclusion that there are no beliefs.[18] Epistemologists face a challenge, then—if one wants to hold onto a viable category of outright belief, one must answer it.

So the situation for belief is in fact rather similar to that for knowledge in the face of skeptical challenges. Although it has received less attention than the knowledge case, a few authors have explored parallel treatments here. One influential example is Brian Weatherson (2005), who invokes a kind of interest-relative invariantism about belief: he argues that although high credence in general does not suffice for belief, that doesn't mean we need maximal credence in order to believe. Instead, on Weatherson's view, the credence that suffices for belief varies according to the subject's practical situation. Another related kind of view is the 'sensitivism' defended in a series of papers by Roger Clarke, according to which credence itself is defined in a shiftier way, with outright belief defined as maximal relative credence.[19] One advantage of this approach over Weatherson's is that it is more general—not only a subject's practical situation, but other factors, like which possibilities are being considered, are relevant for whether a subject's state counts as one of outright belief. Since Clarke's approach to belief is the nearest inspiration for my own, it will be helpful to review it in some small detail. I turn to this task in the next section.

7.3 Clarke, Sensitivism, and Belief as Credence One

Clarke's *sensitivism* is a kind of relevant-alternatives approach to belief. The basic idea is that to believe that p is to be in a doxastic state according to which all relevant not-p possibilities are ruled out. This basic idea is expressed with several technical notions: 'doxastic state', 'relevant', and 'ruled out'.[20] The subject's doxastic state, for Clarke, contains more structure than lists of those propositions the subject believes—instead, it's represented by a kind of probability distribution over all possibilities. Which possibilities are *relevant* for belief will vary according to the subject's situation; it corresponds roughly to the idea that the subject is 'taking a possibility seriously'. (You are probably being reminded of Lewis's 'Rule of Attention'—that's good.) Among those possibilities that are relevant in a given situation, some but not others are *ruled out* by the subject. Clarke says less about what it is to 'rule out' a possibility than one might like, but one might take it as a sort of

[17] One influential such case is Davidson (1963).

[18] There have also been *eliminiativists* about belief—see e.g. Churchland (1981) and Churchland (1986). Although such eliminativists don't think there are any beliefs, they do not reject belief for anything like the reason gestured at here; it's not that they think real-life doxastic commitments aren't strong enough to suffice for belief—they reject the notion of doxastic commitments altogether.

[19] (Clarke, 2013), (Clarke, 2015), (Clarke, 2017)

[20] One might inquire into 'possibilities' too, but we can safely leave that as a placeholder for now; I am not sure what Clarke's own view is about possibilities, but it will be fine to think of them for now as 'possibilities' in the same sense as in my Lewisian approach to knowledge ascriptions; see §1.7.

proto-belief primitive. The idea is intuitive enough: ruling out a possibility is rejecting it; it's taking it to be established not to obtain. (This contrasts with not taking it seriously, whereby one might merely be *presupposing* that it doesn't obtain.) Given this framework, a natural approach to belief has it that, in Clarke's words,

> the believer's context determines a set of relevant possibilities, and to believe p in a context is to rule out all not-p-possibilities relevant in that context, whatever those possibilities might be.[21]

I should note that Clarke is not using the word 'context' in the linguistic sense in which I've been using it throughout this book; his talk of 'the believer's context' is talk of the relevant aspects of the believer's situation. Even though he describes it at times as the view that 'belief is context-sensitive', Clarke's sensitivism makes no semantic claims about 'belief' ascriptions. It is a metaphysical view about the nature of belief, not a semantic view about the truth-conditions of belief-ascriptions.[22] In the next section I will suggest that there is good reason to *prefer* a contextualist implementation of Clarke's basic idea, but my project for now is merely to explain Clarke's framework.

Consider an example. Suppose that Sir Marmaduke, a connoisseur of pedigree, has researched Aline's family tree and discovered that she descends directly from Helen of Troy. He has a clear memory of consulting the various authoritative peerages that record these facts of lineage. And—although this isn't strictly needed to make the point—let's suppose that everything is as it seems. Aline *is* a direct descendant of Helen, and Marmaduke's books accurately and reliably recorded that fact. Relative to modest epistemic standards, Marmaduke *knows* that Aline descends from Helen. Nevertheless, Maramaduke is not an overly dogmatic person; he doesn't think it's *impossible* for someone for whom things seem the same as they do for him to be wrong about someone's ancestry. He knows that there are family secrets about parentage, and that it's possible for the peerages not to reflect genuine ancestry. He thinks this happens only quite rarely—and let's stipulate that he's right—but it's not like his mind is closed to the *possibility* that something might have gone wrong. Marmaduke's doxastic state is, I should think, not at all unrepresentative of most of our doxastic states with respect to most of the things we believe: there is a sense in which he is committed to it—he doesn't merely believe that she was *probably* descended from Helen of Troy; he really believes that she *was*. But there is also a part of him that recognises that he could be wrong. This is precisely the tension that Clarke's view is designed to resolve.

The tension in question is one that obtains with *a single doxastic state*—it's *not* just that sometimes he's confident that she descends from Helen of Troy, and other times, when reflecting on error possibilities, he isn't. This would be a diagnosis along the lines of Jennifer Nagel's 'adaptive invariantism'.[23] Clarke doesn't posit a change in *doxastic state*, though he will posit a change in belief. (I shall return to this feature of Clarke's view in §7.4.) Marmaduke's doxastic state can be represented as in Fig. 7.1.

The rectangle is the space of all possibilities; it is divided into those in which Aline is, and is not, descended from Helen of Troy. The shaded part of the diagram indicates those possibilities that Marmaduke *rejects*; any unshaded possibility is one which is *doxastically possible* for Marmaduke. He has rejected most of the possibilities in which Aline is not descended from Helen, because (he takes it that) his evidence conclusively rules them out. For example, he rejects p_1, a possibility in

[21] Clarke (2015, p. 9). This is a simplification of Clarke's considered approach, which invokes a plausibility ordering for possibilities. Clarke (2015, p. 11) This ordering is important for the extension of his sensitivist framework from outright belief to credence in general—see Clarke (2013). His more complex view is a generalisation of the one stated here; the simpler version includes sufficient machinery for my purposes.

[22] Clarke (2015, p. 8) Clarke's earlier papers are naturally interpreted in an invariantist framework, but Clarke (2017) embeds the thought within a contextualist approach.

[23] Nagel (2010). See fn. 14.

Figure 7.1: The shaded area is that which Marmaduke rules out.

which the ancestry books faithfully record that Aline was *not* descended from Helen. Marmaduke's experience with the books is inconsistent with that possibility. The few possibilities to the right of the line that do remain are those skeptical scenarios about which Marmaduke has a kind of epistemic modesty. He does not consider himself to be in a position to rule out certain elaborate error possibilities—ones like p_2, for instance, in which Aline's great-great-grandfather was the secret son of someone not recorded in the peerage. (Marmaduke also rules out some possibilities in which Aline *was* descended from Helen of Troy—for example, those in which she's descended from Helen, and, moreover, the Trojans defeated the Greeks; his historical education leads him to rule those possibilities out. This is why some of the possibilities to the left of the line, such as p_3, are shaded.) The doxastic state of the subject is in an important sense *prior to* any facts about belief.

On Clarke's approach, belief is a kind of generalisation over certain subsets of the subject's doxastic state. As he puts it, in different "contexts", subjects' doxastic states realise belief in different ways. (I'll use 'situation', continuing to reserve 'context' for a feature of a speaker's linguistic environment.) Different subject situations, according to Clarke, determine different sets of relevant alternatives. When Marmaduke is talking up Aline's many virtues to his son, he is ignoring skeptical scenarios like those in which Aline's ancestry differs from that recorded in the ancestral volumes. But in a different practical or conversational situation—one in which Marmaduke is expressing the kind of epistemic humility we've been discussing, for instance—the doxastic state may remain just the same, but the set of relevant possibilities will broaden. These two sets of possibilities are indicated in Fig. 7.2.

The inner circle corresponds to the less skeptical situation, where Marmaduke is *assuming* that the books are reliable—p_2 is ignored as irrelevant. Since his doxastic state eliminates all not-p possibilities in that circle, he counts as believing that p in this situation. The broader circle, by contrast, represents a more inclusive set of relevant possibilities, including some not-p cases (p_2 for

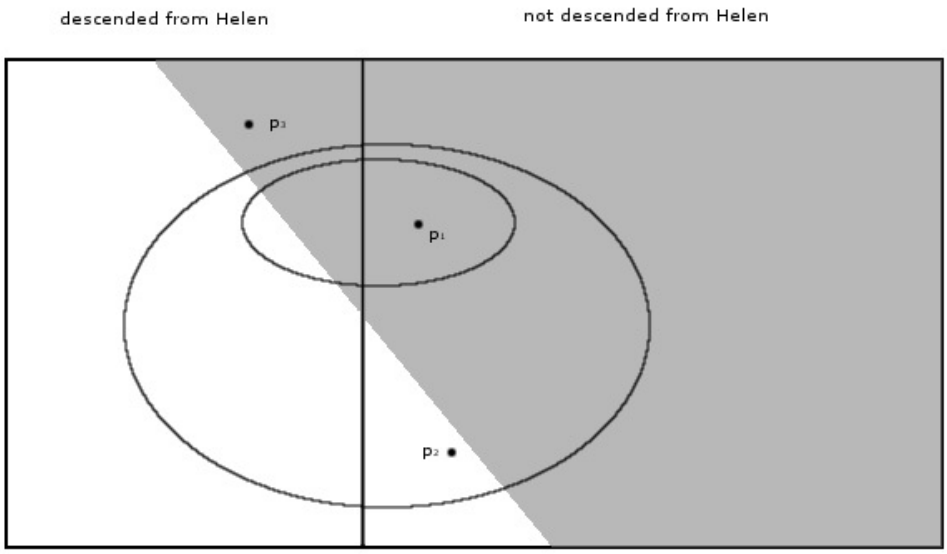

Figure 7.2: Two possible sets of relevant possibilities.

instance) that Marmaduke's doxastic state does not represent as ruled out. Consequently, when these possibilities are considered, the doxastic state does *not* suffice for believing that p, since there are relevant not-p possibilities that aren't ruled out.

There are a number of motivations for an approach to belief along these lines. One is that it may deliver the best solution to belief-based versions of the preface paradox and lottery paradox; another is that it may be the only way to retain attractive closure principles about belief and rationality.[24] But it is also simply an elegant approach to a puzzling question that needs answering, concerning the tension between epistemic modesty and the full commitment of belief. There are however drawbacks to some of the details of Clarke's approach; engaging with some of them will motivate the contextualist implementation that I prefer.

7.4 Challenges for Clarke

One challenge to Clarke's picture concerns whether it really yields the result he claims for it—that a subject's belief state changes from situation to situation, even though her doxastic state itself remains constant. Given the stipulative articulation of what a doxastic state is, it is not difficult for Clarke to make this come out *true* on his view, but there is room to wonder whether Clarke's is really a view according to which what doxastic state it takes to count as believing varies. The point here is related once again to the difficulty of describing just what it takes to be a 'subject-sensitive invariantist' about knowledge. Here is a way of bringing out the issue: Clarke's view is one according to which subjects' beliefs change according to their practical situations, and according to which possibilities they're considering. When Marmaduke is presupposing that the peerage is accurate, he believes

[24] On the preface paradox, see (Clarke, 2015); on lotteries, (Clarke, 2013). On closure, see (Worsnip, 2015).

that Aline descends from Helen; when he's thinking about the possibility that there may be an error, he stops having that belief. Clarke's view is one according to which one's beliefs change, depending on what one is thinking about. And indeed, as Worsnip (2015, p. 7) observes, it is not only one's outright beliefs that, on Clarke's view, change according to the subject's situation—it is the subject's credence too.[25] Since on Clarke's view, subjects stop believing when stakes go up or error possibilities are mentioned, and their credences also go down, the approach has quite a lot in common with that of theorists like Jennifer Nagel, who explain what happens in such cases in terms of a psychological weakening of confidence.[26] Although Clarke's theory does have a stipulated 'doxastic state' that remains constant—that which is represented by the shaded area in the diagrams above—one wonders how substantive this constancy is. After all, Clarke's view has it that which possibilities are relevant is a matter of the subject's individual psychology. So there does seem to be a straightforward and important sense in which the way the subject is thinking about the question is influencing which beliefs she has.

A closely related objection, pressed by Worsnip, is that it doesn't in general seem rational for subjects to change their beliefs or credences when the stakes go up or more error possibilities are brought to salience; but on Clarke's view, they typically will. Furthermore, not only does it not seem *rational*, it doesn't seem like, in at least many such cases, this *is* something we're doing. Worsnip writes:

> [A] credence, as I understand it, is (or at least commits one to) a kind of estimate of the probability or likelihood of a proposition. Consequently, when one changes one's credence for some proposition, one is thereby committed to thinking that *either* the probability of the proposition has changed *or* that one made a mistake in one's previous estimate of the probability. But in [cases like these], neither option seems plausible. It is very unnatural for me to think that, after I have been offered the bet, the probability … has changed. But nor does it seem like, after I am offered the bet, I need think that I made a *mistake* in my original assessment of the probability. Worsnip (2015, pp. 7–8)

This kind of objection is very similar to many of the standard objections to so-called 'subject-sensitive invariantism' or 'interest-relative invariantism' about knowledge.[27] Against the view that raising the practical stakes makes it more difficult to have knowledge, even if it doesn't influence any other epistemic factors,[28] one may object that it's implausible to suppose that, for instance, a subject in a low stakes situation has knowledge, but *wouldn't* have knowledge if the question were more important. Or similarly, if the question is important today but wasn't important yesterday, it sounds somewhat bizarre to suppose that the subject knew yesterday but doesn't any more, when

[25] As indicated in the previous section, I am working with a simplified version of Clarke's approach that does not generalise to credences. On Clarke's own view, instead of merely ruling out possibilities or not, subjects assign a level of 'plausibility' to possibilities, and credence is defined as a probabilistic measure over those plausibilities, given a set of relevant alternatives. Since Clarke treats outright belief as credence one, relative to the set of relevant possibilities, any positive plausibility assigned to a possibility suffices for not being ruled out.

[26] See the discussion of Rubin (2015) on the commitments and plausibility of pragmatic encroachment on rational credence.

[27] See e.g. Russell and Doris (2008), Schaffer (2006), and (DeRose, 2009, Ch. 6).

[28] Spelling out just what remains the same is vexed, particularly within a knowledge-first framework. One can't stipulate that the *evidence* is the same, if one subscribes to E=K. The stipulation that the *level of confidence* is the same is a more bizarre one than epistemologists sometimes realise, given psychological regularities about how humans respond to high-stakes situations—and even if one did, then arguably, one couldn't stipulate that one's epistemic responsibility remains the same. See again Nagel (2010) and Nagel (2011). Something like this complication may be part of why Jason Stanley suggests (though he doesn't defend it in his book) that *all* epistemic states are stakes-sensitive. Stanley (2005, p. 124) (But see Rubin (2015) and Ichikawa et al. (2012).) It is also part of the reason it's quite difficult to come to grips with just what the content of 'interest-relative invariantism' is. See §1.9.

the only change is an increase in the stakes. A directly parallel objection to Clarke's view about belief suggests would ask, how plausible is it that, although a subject's doxastic state stays the same, she believed that p yesterday but doesn't believe that p today? Or that she believes that p, but, if the question were more important to her, she wouldn't have believed that p, even though she would have been just as confident?

Proponents of this kind of 'shifty invariantism' have various strategies in response to this kind of challenge; typically, they will invoke a kind of error theory whereby one mistakenly project's one's current or actual situation onto the past, future, or hypothetical one. But a contextualist implementation of the basic idea can accommodate the same kind of consideration without such an error theory.[29]

7.5 Contextualism about Belief Ascriptions

The contextualist version of Clarke's view would retain Clarke's basic framework, whereby believing that p is a matter of ruling out all not-p possibilities within some restricted set of relevant possibilities, but unlike Clarke's version, the contextualist would have the set of relevant possibilities determined in part by the conversational context of the speaker making the belief ascription. This would allow us more cleanly to separate the roles of doxastic states—now a more comprehensive description of everything relevant about the subject—and the set of relevant possibilities, which can vary more freely with the conversational interests of those speaking. The 'belief' contextualist draws something very much like the distinction Clarke's view made—one and the same doxastic state can suffice for 'belief' relative to some standards, but not to others—but it does so in a way that does not require any change in the subject's psychology at all.

Consider again Sir Marmaduke after consulting his peerage, and learning about Aline's noble ancestry. As before, we can represent his doxastic state as one in which most possibilities where she isn't descended from Helen as ruled out, leaving only a small number of far-fetched possibilities about which he's agnostic. And as before, given different sets of relevant possibilities, we can consider Marmaduke to 'believe' that she's descended from Helen relative to some weaker standards, but not to 'believe' relative to a stronger standard. But on the contextualist approach, we don't need those standards to correspond to differences in the subject's attention or practical situation; different conversational contexts can generate those standards, even holding the subject's situation fixed. One might expect this to work in a way very similar to the contextualist dynamics that I've suggested affect 'knows'. To get a sense for why one might prefer an approach like this, consider these two conversational contexts; let Marmaduke's doxastic state be the same as in the previous discussion—that represented by fig. 7.1 (the diagram that did *not* include specifications of the sets of relevant possibilities). Since outright belief is a technical philosophical notion, the conversations we will consider are taking place in epistemological contexts. So *unlike* in the case of the dialogues motivating contextualism about 'knows'—which *is* a colloquial term—I will not make a particular attempt to make these dialogues sound colloquial. However, I hope they are recognisable as the sorts of things *philosophers* might naturally say. This is the appropriate methodology as I'm offering a treatment of this *philosophical* bit of language.

Lower Standard:

Mrs. Partlett: Why does Sir Marmaduke want his son to marry Aline?

[29] This point has been contested— Stanley (2005, pp. 116), for example, argues that contextualists need their own error theory in order to accommodate the relevance of the subject's practical situation. But this relies on a a particular conception of epistemic standards that I do not endorse. See my §1.8. There's no reason subjects' practical interests can't have an effect on speakers' epistemic standards. See DeRose (2009, Ch. 7) and Fantl and McGrath (2009, Ch. 2).

Dr. Daly: Noble lineage is important to Sir Marmaduke; he desires that his family be joined with another family of strong descent. And, having consulted the relevant peerages, he has found that Aline is descended from Helen of Troy.

Mrs. Partlett: And does he trust the peerage and believe what it says, or does he take a skeptical attitude about its contents?

Dr. Daly: He trusts it. Marmaduke believes that Aline is descended from Helen.

Higher Standard:

Hercules: Sir Marmaduke has read in the peerage that Aline is descended from Helen of Troy.

Mr. Wells: Oh, and is Sir Marmaduke certain that peerages never contain mistakes?

Hercules: No, he knows that they're occasionally inaccurate. But he's not dwelling on skeptical scenarios or letting them interfere unduly with his planning. He's not even thinking about far-fetched possibilities; he's just assuming that the peerage is right in this instance.

Mr. Wells: Ah, so he's just *presupposing* that the books are correct; he doesn't represent himself as having actually *established* that.

Hercules: Right.

Mr. Wells: So he doesn't (outright) believe that Aline is descended from Helen; that depends on assumptions that he knows full well that he hasn't demonstrated.

On the contextualist approach, the lower-standard context generated in Mrs. Partlett's conversation sets the set of relevant possibilities more modestly, and including none of the not-p possibilities that Marmaduke doesn't reject; in Mr. Wells's higher-standards context, some such possibilities are included, and so Marmaduke doesn't count as (outright) believing. But on the contextualist approach, unlike Clarke's approach, this difference can exist without any psychological difference in Marmaduke whatsoever—as in the case of divergent standards in knowledge ascriptions, we may imagine that the two conversational contexts are taking place in the same world and at the same time, about the same individual. So the contextualist approach has a clearer explanation for why we should think of the view as one in which the *standards* for belief change; it represents a starker contrast than Clarke's does with views like Nagel's according to which subjects sometimes lose confidence in the face of high stakes or error possibilities, and thereby stop believing.

The approach is also well-positioned to avoid Worsnip's objection to Clarke mentioned above. Since nothing changes about the subject between these two conversational contexts—he doesn't stop believing; he merely counts as 'believing' relative to one context but not to another—there's no reason we should suppose that, from his own point of view, he should represent himself as having previously made a mistake. Similarly, the problem for modal and temporal embeddings is dissolved, in a way exactly analogous to the contextualist treatment of the similar issues about 'knows'. Assuming that Marmaduke's doxastic state hasn't changed, a sentence like 'yesterday he believed that p, but today he doesn't', when the only change that's occurred is one in the speaker's conversational context, will come out false; whether 'yesterday he believed that p' is true depends on the speaker's current context—if it sets a demanding standard for belief, it does so for both conjuncts.

This diagnosis depends on Marmaduke's doxastic attitudes. So far we've been stipulating that he's open-minded about the relevant skeptical possibilities. But it's possible for him to be more dogmatic with respect to Aline's ancestry. There is a possible version of Marmaduke who doesn't merely *presuppose* the accuracy of the peerage, but who affirmatively *rejects* the possibility that

the peerage is inaccurate. When he himself is engaging with relatively low standards, the difference won't be much apparent—the difference between someone who ignores a counterpossibility because it's a far-fetched scenario not to be bothered thinking about, and someone who ignores it because he thinks he knows it to be false will only occasionally manifest in behaviour. But the differences will be evident when he's *asked* about those scenarios, or when the stakes go up dramatically. Suppose that a *dogmatic* version of Marmaduke thinks that Aline descends from Helen, and furthermore takes himself to be in a position to rule out the possibility that the peerage is in error in this instance. He will not admit that he's ultimately relying on the assumption that the book isn't wrong in this instance—he'll take himself to be in a position conclusively to rule that out. He may also manifest his higher confidence in betting behaviour—our original, more epistemically modest version won't bet on her ancestry at terrifically long odds, but the more dogmatic one might. His doxastic state might be represented as in fig.7.3.

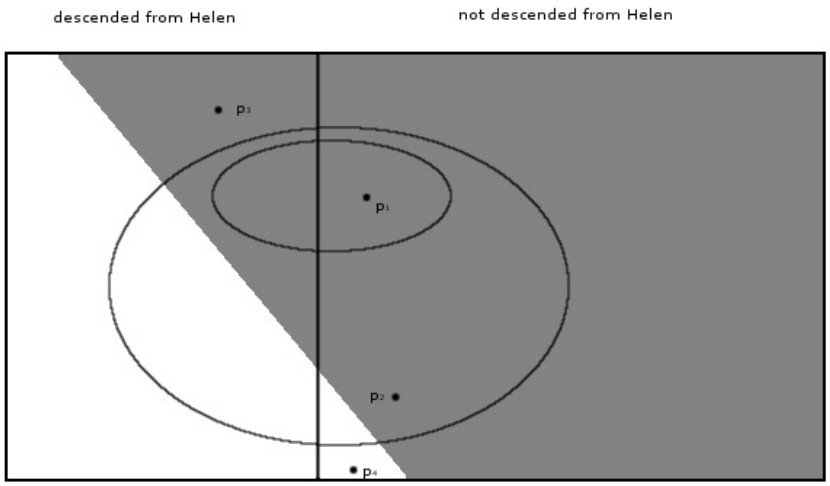

Figure 7.3: The shaded area is that which Marmaduke rules out; it now includes p_2

In this version, p_2, the possibility that the peerage is in error, is actually shaded—so even relative to contexts where that possibility is considered relevant, Marmaduke still counts as 'believing' that Aline is descended from Helen. I have left some even further-fetched not-*p* possibilities like p_4 unshaded—perhaps, for instance, even this version of Marmaduke is unwilling to take himself to be able to rule out the possibility that the universe was created only last year, complete with fabricated memories and a matching historical record. So relative to *very* skeptical contexts, he might not count as 'believing' that Aline descends from Helen, although he still has that belief relative to both the more and the less skeptical contexts marked on the diagram. But one could stipulate an *even more* dogmatic version who rules these possibilities out as well. So the framework posits a significant distinction between that which is *presupposed* and that which is *ruled out*.

With the framework in place, we can now return to the normative connections about the relationship between 'believes' and 'knows' in particular contexts.

7.6 'Knows' and 'Believes'

I've now set out an approach to both 'knows *p*' and to 'believes *p*' that makes each centrally require, of some set of relevant possibilities, some kind of elimination of 'all' the not-*p* possibilities within that set. Two obvious questions arise: what is the relationship between those sets of possibilities, and what is the relationship between those kinds of elimination?

Let's start with the latter question, which may be more straightforward. The 'ruling out' of possibilities, in the sense relevant for whether a subject has an outright belief, seems closely tied to the notion of evidence. Indeed, I've been employing the Lewisian language of evidence 'ruling out' possibilities throughout this book. A plausible normative principle suggests itself: a subject should only reject a possibility if her evidence is inconsistent with it. On the view of Ch. 3, 'evidence' is also a context-sensitive term, so to speak more precisely, the principle in question is that a subject should only reject a possibility *s* if, relative to a context in which '*s*' is relevant, the subject's 'evidence' is inconsistent with *s*'s obtaining. Equivalently, one should only reject those possibilities that are inconsistent with one's *basic* evidence, in the sense of §3.10. This is only a necessary condition for the rejection of possibilities—it's not sufficient because there can be cases where a subject is unable to tell that a possibility is inconsistent with the evidence; in such cases, one is not in a position to have a belief that is properly based on the right kind of evidence. See fn. 31 in §1.7.

In the case we've been considering, it seems rather plausible that Marmaduke's evidence does *not* eliminate the possibility that the peerage is mistaken about Aline's ancestry. None of his basic knowledge seems to rule out scenarios, for example, whereby Aline's great-great-grandfather was secretly parented by someone other than those recorded in the books. Assuming that this is correct, then the dogmatic version of Marmaduke considered in the previous section is violating a doxastic norm—he is ruling out possibilities that are consistent with his evidence. The norm in question commends the original version we considered, whose doxastic state left open possibilities like p_2; on that version, he counted as 'believing' that Aline descends from Helen relative to lower-standards contexts, but not relative to higher-standards contexts. (He also counted as 'knowing' this fact relative to low-standards contexts, and not relative to higher-standards contexts. More on this connection shortly.) One ought not to *reject* possibilities consistent with one's evidence, although it is perfectly fine to *ignore* them and *presuppose* them false.

It is less obvious what to say about more radical skeptical scenarios, like the possibility that Marmaduke is a brain in a vat. If his basic knowledge includes external-world objects of factive perceptual states, then his evidence does rule those possibilities out—if not, it probably doesn't. So the question of whether a subject ought to reject, or merely to ignore, BIV possibilities turns on this substantive question about basic knowledge. As in the previous discussion, I won't try to settle that question here, although I do have some sympathies with the richer conception of evidence, corresponding to a more Moorean reaction to these cases.

What of the relationship between the relevant sets of possibilities for belief ascriptions and knowledge ascriptions? It would be tidy to identify them, as I tried to identify those for knowledge ascriptions and counterfactual conditionals in Ch. 2. Unfortunately, this tidy answer won't do. Knowledge ascriptions (and counterfactuals) always require the subject's *actual* world to be a relevant possibility. If the actual world could be irrelevant for the truth of a knowledge ascription, factivity could fail; if it could be irrelevant for a counterfactual, modus ponens could fail. But it does not seem to be a requirement for belief ascriptions that the actual world be relevant. The easiest way to appreciate this is to think of cases in which something surprising—the sort of thing one might naturally presuppose false—turns out to be true. Suppose for instance that the Notary is considering having a second cup of tea. He thinks he'd enjoy it, but he thinks it would prevent him from sleeping tonight. The Notary believes that the tea contains caffeine—he rejects all relevant possibilities in which it doesn't. And he is deliberately open to the possibility that he will drink the tea and have trouble

sleeping as a result. One thing he's confident about, though, is that the tea certainly wouldn't *help* him to sleep. Tea never has that effect on him. He's never even considered the far-fetched possibility that the tea contains a magic love potion that will put him instantly to sleep. And indeed, he has no reason whatsoever to consider that possibility. Though let's suppose that he doesn't *reject* it—he merely presupposes that it's false. Relative to his own conversational context, and that of anyone else who is ignoring that bizarre possibility, we can represent the Notary's doxastic state and belief as in Fig. 7.4.

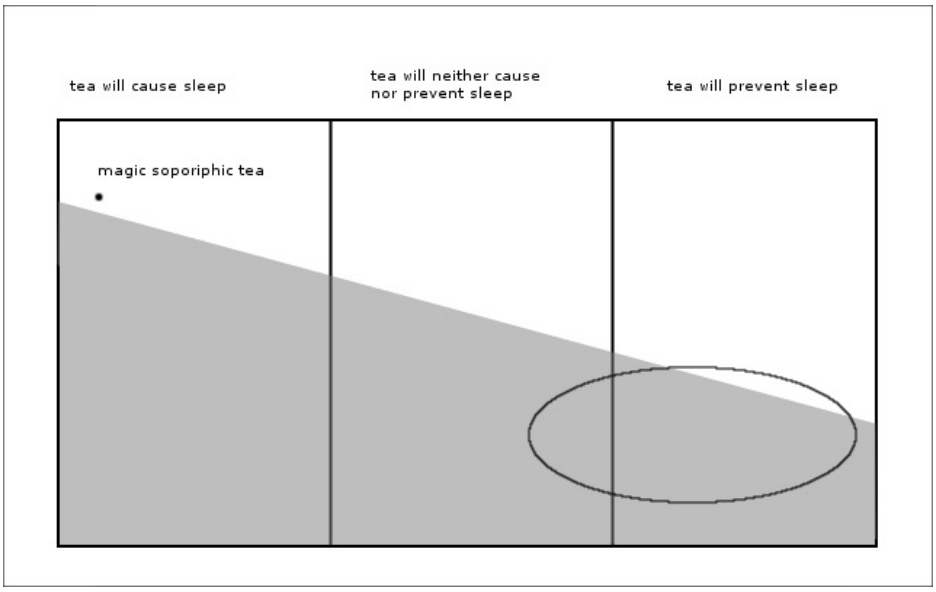

Figure 7.4: The notary believes that the tea will hurt his sleep; he presupposes that the magical possibility does not obtain.

This can be an accurate description of the situation, even if it turns out that as a matter of surprising fact, the magical possibility is actual; there really is a magic philtre in the tea, and it will put the Notary to sleep (perhaps also making him fall in love). The fact that it's an error to presuppose something just doesn't interfere with it's being genuinely presupposed; actual possibilities can be irrelevant for the purpose of belief ascriptions, even though they can't for knowledge ascriptions. (If they could, then the Notary could be truly said, relative to those contexts that ignored the actual surprising truth, to 'know' that the tea won't make it easier for him to sleep. This is why we keep an actuality constraint on knowledge-relevance.) Relevance for belief ascriptions is in this respect more similar to Stalnakerian contexts than to relevance for knowledge ascriptions.[30]

But although the sets of relevant possibilities aren't identical, it remains plausible that there are systematic connections between them. For example, it seems quite plausible that all belief-relevant possibilities are knowledge-relevant. Without this relationship, it's difficult to see how 'knows' could entail 'believes' in every context. (Suppose that in a context, possibility p is belief-relevant, but not knowledge-relevant; then if p is not ruled out, but rather merely presupposed against, then it falsifies 'believes p'. But it seems the subject might still satisfy 'knows p', if, within the more limited set of knowledge-relevant possibilities, the subject's evidence rules out 'all' not-p possibilities, and the subject responds in the appropriate way.) I think there's also something to the line of thought that

[30]Clarke (2016) compares belief and assertion in a related way.

says that a belief-relevant possibility is one that is being treated as a live possibility, in a way that, given something like the rule of attention, will suffice for knowledge-relevance.[31] So let's tentatively posit this connection:

(1) All belief-relevant possibilities are knowledge-relevant possibilities.

Another connection—a normative one—between the approach to belief and that to knowledge would relate rejection to elimination by evidence:

(2) It is correct for S to reject a possibility if and only if S's evidence is inconsistent with that possibility.

There are two directions to this principle. The left-to-right direction has it that all cases of correct rejection are cases of elimination by evidence; it is incorrect to reject a possibility that is consistent with the evidence. This is very plausible—it encodes the idea that we should not let our doxastic attitudes go beyond what is warranted by our evidence. The right-to-left direction has it that any time the evidence eliminates a possibility, it is correct to reject it. This direction may be more controversial; one may question it on the grounds that sometimes, evidence is inconsistent with a possibility, but the subject hasn't noticed that fact, or perhaps even lacks the cognitive resources to appreciate it. Notice, however, that the principle doesn't say that any time one rejects a possibility inconsistent with the evidence, one has done well—it says only that one has done what it is correct for one to do. We can still recognise the difference between doing that which is correct, and doing that which is correct *because* it is correct. (Compare what it is for an answer on a mathematics exam to be 'correct'.)

While I don't consider either (1) or (2) to be *obvious*, they seem to me to be attractive hypotheses. A third assumption I wish to consider is this:

(3) Any possibility consistent with the subject's presuppositions is belief-relevant.

Notice that this is a substantive assumption, going beyond anything I've posited so far. The contextualist approach to 'belief' has it that the presuppositions *of the speakers* are relevant for which possibilities are relevant; (3) has it that those *of the subject* are relevant too. There is no contradiction in accepting the framework and denying (3). However, we shouldn't be too surprised at the suggestion that both speaker and subject have roles to play at this level. After all, it is clear that in the case of the Lewisian approach to 'knows', the speakers and the subject each have roles to play in determining which possibilities are relevant. As I've emphasised a number of times, we can't think of the speakers' conversational context as being sufficient to establish the set of relevant possibilities. *One* reason this is so is the factivity of 'knows', encoded in the rule of actuality—this one doesn't apply to belief; but we shouldn't be too surprised if it turns out there are others.

Furthermore, I think, on consideration of cases, that invoking (3) gives us the right result. Consider what a counterexample to (3) would look like. It'd be a case in which, although the subject herself treats a certain kind of skeptical not-p possibility as live, this fact is irrelevant for whether or not she satisfies 'believes p'. Suppose for instance that Aline's mother is planning a wedding, since she knows that Aline and Alexis are engaged to be married. Suppose further that, with respect to the proposition that Aline and Alexis *will* be married, her doxastic state is as indicated in fig. 7.5. She has a moderate attitude towards that proposition—she rejects all the commonplace possibilities

[31] As indicated in §1.7, I think the rule of attention is only approximately correct, but the kind of 'treating as live' is much closer to what is needed.

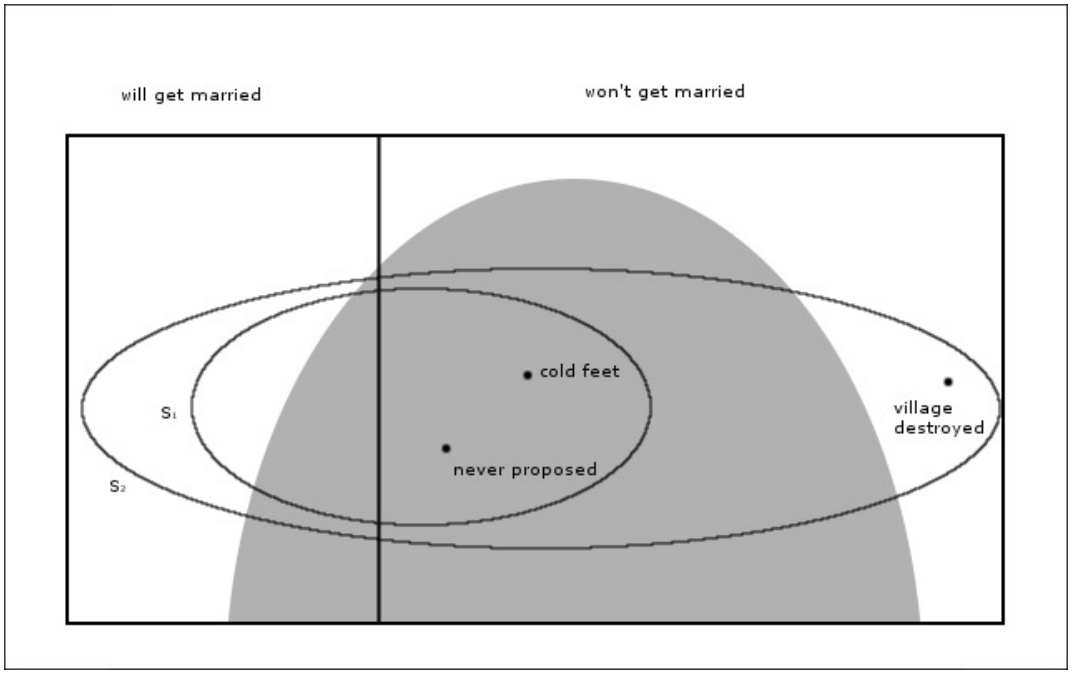

Figure 7.5: S_2, but not S_1, includes unrejected possibilities in which the wedding won't occur.

where the marriage won't occur, even while leaving open certain far-fetched possibilities, such as those in which the village is suddenly and unexpectedly destroyed before the wedding.

As we've seen in the discussion of the parallel case of Sir Marmaduke's belief about Aline's ancestry, the contextualist approach to 'belief' allows that this state suffices for 'belief' in some contexts—those in which any such possibilities are presupposed to be false—but doesn't in others—those in which they're treated as live. Two sets of possibilities corresponding to such contexts, S_1 and S_2 respectively, are also indicated in fig. 7.5. (For each set of possibilities, the relevant question is whether there is any white space to the right of the line.) However, suppose we add an additional stipulation to the case: suppose we stipulate that the subject *herself* is not presupposing that life in the village won't suddenly end soon. The two sets of possibilities are not on a a par; the subject herself is making the presuppositions corresponding to S_2. This stipulation should not be thought of as changing which possibilities she *rejects*—nor even as changing her credences or her estimations of probabilities. Just like the possible version of her who presupposes that this won't happen, she considers it *extremely unlikely* that all the villagers will suddenly drop dead within the next week. There is a straightforward sense, then, in which her *doxastic* state is no different than it'd be if the presupposition were in place. Nevertheless, it is a fact about the subject—very plausibly a relevant one—that she is treating certain skeptical possibilities as live options. If you ask her, 'will Aline and Alexis get married in the future?', she won't give an unqualified 'yes'; rather, she'll say something like: 'assuming something very unexpected, like the sudden death of everyone in the village, doesn't happen, yes, they will'. It seems to me that this psychological fact about the subject is relevant for whether she satisfies 'believes that Aline and Alexis will get married', even in contexts in which the speakers are making stronger presuppositions. My suggestion is that it is a constraint on belief-relevant sets of possibilities that anything treated as open *by the subject* is relevant, in the same way that anything treated as open *by the conversational participants* is also relevant. In other words, I

suggest that (3) is true.

So there is an additional feature of the subject's psychology, beyond the distribution of plausibility to possibilities, that plays a role here. In some situations, it may be useful to encode this feature in the diagrams I've been using to characterise the subject's doxastic state; the case of Aline's mother who is being consciously open-minded about skeptical possibilities might generate a diagram like fig. 7.6.

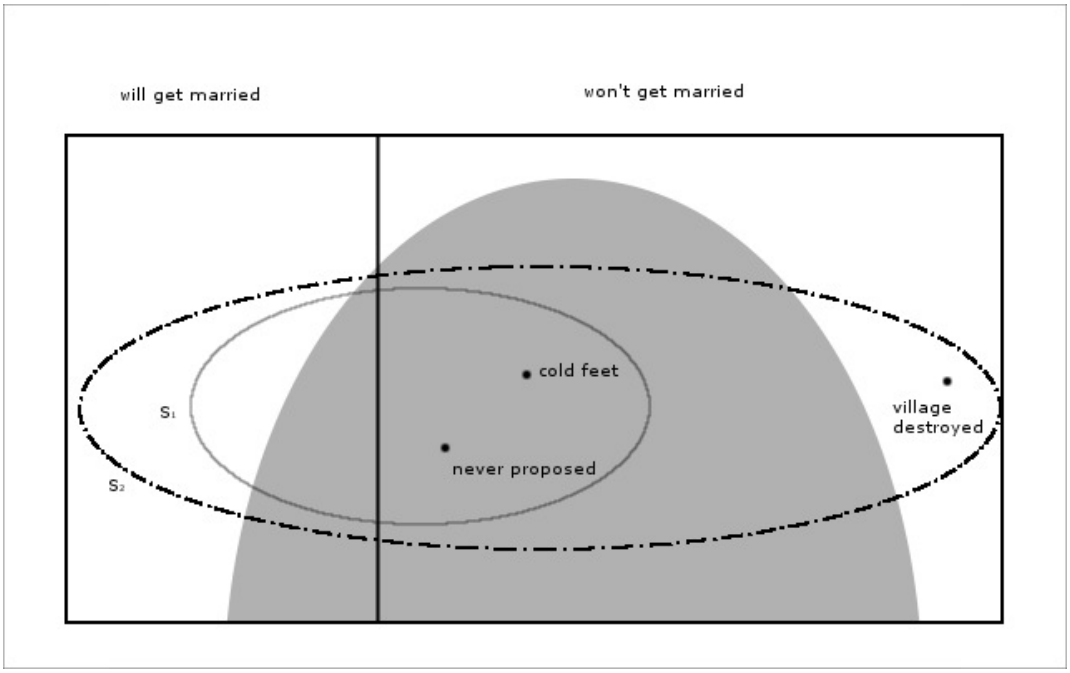

Figure 7.6: The dashed circle represents the *subject's* presuppositions, which play a privileged role.

Here, the added structure shows what the subject herself is treating as live; (3) would have it that it is a constraint on the sets of belief-relevant possibilities that, regardless of the context, the set must at least include everything inside the dashed region. So even relative to a context in which the skeptical possibilities are presupposed false, 'She believes that Aline and Alexis will get married' will be false, because she does not reject some of the relevant not-*p* possibilities; they are relevant because *she* treats them so, not because the conversational participants do. Whether one wants to describe this as part of the subject's *doxastic* state represents a terminological choice point. On the one hand, it really is a feature of the subject that makes for a difference in whether she believes—not merely in whether 'believes' truly applies to her—so there's a case to be made that it should be thought of as doxastic.[32] On the other hand, the simpler picture that does not include the subject's presuppositions does give a full story of what the subject represents her evidential state to be, and so one might prefer to consider such pragmatic features as these not to be strictly 'doxastic'. The terminological choice doesn't matter; the important thing is that a subject has two roles to play in questions about whether 'believes' applies: the subject's presuppositions can make a difference for which possibilities are relevant, and the subject's ruling-out attitudes make a difference for whether she satisfies 'believes', given a set of relevant possibilities.

[32]Perhaps this is even the best way to think about Jennifer Nagel's treatment of 'epistemic anxiety'—dwelling on certain possibilities and treating them as live partly constitutes failing to believe. See Nagel (2010).

Again, although the resultant picture is slightly more complex than one might have thought, this does not represent any kind change in principle from the relevant alternatives approach to knowledge ascriptions I've been considering throughout the book. According to orthodox Lewisian contextualism, features of the context *and* features of the index, including the subject's psychology, are relevant for determining which possibilities are relevant. To a reasonable approximation, one might even *identify* the set of belief-relevant possibilities with the intersection of the set of possibilities compatible with the presuppositions of the conversational context with the set of the possibilities compatible with the presuppositions of the subject. (Or equivalently: with the set of possibilities compatible with what is presupposed by both the speakers and the subject.) But the truth of (3), along with (1) and (2), is the important part for the attempt to explain and vindicate the knowledge norm of belief; I turn to that project now.

7.7 Knowledge and Proper Belief

I wish to derive a substantive normative connection between knowledge and belief, using the contextualist framework given for each, making use of plausible assumptions. Again, those assumptions are:

> (1) All belief-relevant possibilities are knowledge-relevant possibilities.
>
> (2) It is correct for S to reject a possibility if and only if S's evidence is inconsistent with that possibility.
>
> (3) Any possibility consistent with the subject's presuppositions is belief-relevant.

The first two assumptions imply that 'knows p' in a context entails that one appropriately 'believes p' in that context: Suppose that S knows that p. Then S's evidence eliminates all the knowledge-relevant possibilities in which not-p; so by (1) S's evidence eliminates all the belief-relevant possibilities in which not-p. So by (2) it is correct for S to reject all of those possibilities; so S's belief is appropriate. Knowledge relative to a context entails appropriate belief relative to that context; we have established one direction of a metalinguistic version of the knowledge norm of belief.

What of the other direction? Can we show that lack of knowledge is inconsistent with appropriate belief?

According to the account given in Ch. 1, S knows p if and only if S believes p, basing that belief appropriately on evidence that eliminates all not-p cases relevant in the context. Suppose that S does not know that p. Then either there are some knowledge-relevant possibilities uneliminated by S's evidence in which not-p—i.e., S doesn't have enough evidence—or S fails to form the belief that p properly based on that evidence. If the latter, then there is clearly no appropriate belief. So consider the former case: there are knowledge-relevant not-p possibilities uneliminated by S's evidence. Either some of those possibilities are belief-relevant, or none of them are. If some of them are, then those are belief-relevant not-p possibilities that are uneliminated by the subject's evidence; by (2) above, they are possibilities the subject ought not to reject, and so, either they are not rejected (in which case the subject doesn't 'believe') or they are inappropriately rejected (in which case the doxastic state is defective). So the remaining cases we need think about are those in which the evidence leaves open some knowledge-relevant not-p possibilities, but where none of these are belief-relevant. We've already shown that no other case could undermine the metalinguistic knowledge norm for belief.

It certainly appears that there can be such cases, given the factivity of 'knows'. Suppose for example that a subject is in a far-fetched skeptical scenario, as in the case of the Notary and the magic tea. Then, assuming that the magical possibilities are being presupposed against, they are

not belief-relevant, but they are knowledge-relevant. Something similar holds for Gettier cases as well. Keep the Notary's doxastic state as before, but suppose now that there's no potion in the tea, but there was a secret plot to put it there that was thwarted at the last minute. (Maybe the tea in the five neighbouring villages contains magic potion.) Plausibly, even relative to many conversational contexts in which possibilities involving magic are *ignored*, a sentence like 'the Notary knows that the tea won't make him sleepy' is *false*, because the possibility that this tea is magic, though not actual, is knowledge-relevant. (Lewis's explanation for this is that it is *saliently similar* to the actual possibility that there is a love potion in the tea the next town over.[33])

So, suppose there are, relative to a context, some knowledge-relevant not-p possibilities that are uneliminated by the subject's evidence. We can show, given the assumptions above, that the subject does not appropriately believe that p, relative to that same context. We do this by cases. If the subject *rejects* the possibilities in question, even though they are consistent with her evidence, then she violates (2), and so does not respond appropriately. So consider cases in which she does not reject them, but rather regards them as not having been ruled out. Next we ask whether they are consistent with her *presuppositions*—although she doesn't reject them, does she ignore them? If she doesn't, then by (3), they're unrejected belief-relevant possibilities in which not-p; so she doesn't satisfy 'believes p'. So consider the case where she *does* presuppose against the not-p possibilities.

Now the key question is this: *why is it* that these possibilities are knowledge-relevant? We've seen one way this could obtain already: because they are actual—the rule of actuality includes them as relevant. Another possible way is because they are *similar to* actuality—the rules of actuality and resemblance make them relevant. But if they are actual, then the subject is mistakenly taking something for granted, and so exhibits a doxastic failure.[34] And if they are not, it must be because they are in some way *similar* to a knowledge-relevant possibility. We've seen no other way for possibilities to be knowledge-relevant. But for the subject to presuppose against one but not the other of these relevantly similar possibilities would be irrationally arbitrary. (Imagine considering it an open question whether the first and second of three identically-appearing structures is a fake barn, but taking it for granted that the third one isn't.) So in these cases too, there will be a defect in the subject's doxastic state.

If this is right, then normative constraints on doxastic statuses, combined with the relationships between knowledge-relevance and belief-relevance, can support a fully metalinguistically general version of the knowledge norm of belief.

7.8 Doxastic States and Epistemology

On the picture of belief I've sketched the rejection of particular possibilities is more fundamental than a subject's beliefs. This is very much analogous to the way that basic knowledge may play a more fundamental role in my framework than does knowledge itself. As David Lewis remarked about the latter case, use of the category 'knowledge' is useful at least in part because it allows us

[33] Lewis (1996, pp. 557–8)

[34] Here is an objection: if there is no reason for a subject to suspect that a possibility is actual, then it is no doxastic failure to take for granted its falsity. For example, a subject should not be criticised for presupposing that someone's testimony is reliable, if she has no particular reason to doubt it. While this objection is an important one, it is an objection to the knowledge norm of belief in general, not to my contextualist implementation of it in particular; cases of apparently appropriate belief that fall short of knowledge are often used in criticism of the idea that knowledge is a norm for belief. My aim is not to convince the knowledge-first skeptic in the truth of the knowledge norm for belief; it is to demonstrate that it fits comfortably into my contextualist framework. So in this dialectical context it is reasonable to rely on (admittedly controversially) externalist intuitions like the intuition that one always fails in a genuinely normative way if one takes something false for granted. Defenders of the knowledge norm typically describe these as cases that are normatively defective, but nevertheless blameless. For a knowledge-first theorist like myself (and one who, unlike e.g. Sutton (2007) and Littlejohn (2014), recognises a separate justification norm, one can also describe these as cases that satisfy some genuine knowledge norms but not others. See Ch. 4.

to approximate our epistemic positions, abstracting away from some unimportant details. The same might well go for belief ascriptions. There is an important sense in which the subject's detailed doxastic state—the information encoded by diagrams of the sort included in this chapter—are more fundamental; they're where the real doxastic action is. Nevertheless, belief ascriptions of the sort described in this chapter can encode that information in useful and non-arbitrary ways. Certainly this is true with respect to epistemology—'belief' gestures at what is easily understood as a kind of approximation of, or an attempt at, knowledge. I expect that the notion of 'belief' in question has roles to play elsewhere, too—in action explanation, for example, or moral psychology. But such an exploration is well beyond the current scope of this project.

Returning to Lewis and the case of knowledge, Lewis seems to think that its cognitive efficiency is the *only* virtue of 'knowledge'-talk—it is a kind of pragmatically useful shorthand. Lewis writes:

> What is it all for? Why have a notion of knowledge that works in the way I described? (Not a compulsory question. Enough to observe that we do have it.) But I venture the guess that it is one of the messy short-cuts—like satisficing, like having indeterminate degrees of belief—that we resort to because we are not smart enough to live up to really high, perfectly Bayesian, standards of rationality. You cannot maintain a record of exactly which possibilities you have eliminated so far, much as you might like to. It is easier to keep track of which possibilities you have eliminated if you—Psst!—ignore many of all the possibilities there are. And besides, it is easier to list some of the propositions that are true in *all* the uneliminated, unignored possibilities than it is to find propositions that are true in *all and only* the uneliminated, unignored possibilities.
>
> If you doubt that the word 'know' bears any real load in science or in metaphysics, I partly agree. The serious business of science has to do not with knowledge *per se*; but rather, with the elimination of possibilities through the evidence of perception, memory, etc., and with the changes that one's belief system would (or might or should) undergo under the impact of such eliminations. Ascriptions of knowledge to yourself or others are a very sloppy way of conveying very incomplete information about the elimination of possibilities. . . .
>
> The only excuse for giving information about what really matters in such a sloppy way is that at least it is easy and quick! But it *is* easy and quick; whereas giving full and precise information about which possibilities have been eliminated seems to be extremely difficult, as witness the futile search for a 'pure observation language'. If I am right about how ascriptions of knowledge work, they are a handy but humble approximation. (Lewis, 1996, p. 563)

If anything like the contextualist knowledge-first approach of this book is correct, Lewis is far too quick in selling knowledge far too short. The ways in which knowledge ascriptions encode more detailed information are neither sloppy nor arbitrary—there is a reason 'knows' is a linguistic universal. Furthermore, the proper explanation of the shiftiness of knowledge ascriptions yields further theoretically significant notions—key among them have been epistemic standards and sets of relevant alternatives. I've tried to argue throughout this book that these ideas too are theoretically useful, in a way independently of their roles in fixing the truth-conditions of knowledge ascriptions; this is further reason to think that the contextualist approach is not ad hoc. For example, if my project in Ch. 2 is correct, the epistemic standards and sets of relevant possibilities invoked by the contextualist are also necessary for a proper understanding of counterfactual conditionals; the incremental knowledge norm of assertion given in Ch. 6 suggests that tracking sets of relevant alternatives plays important roles in regulating assertion, in addition to their roles in theorising

about knowledge. While capturing a surprising pattern of intuitions about knowledge ascriptions has always been among the central motivations for contextualism, I think it's at least as important a mark in contextualism's favour that it can also play a central role in clarifying and explaining the theoretical significance of knowledge.

Bibliography

Anderson, Charity. 2014. "Fallibilism and the Flexibility of Epistemic Modals." *Philosophical Studies* 167:597–606.

Annis, DB. 1978. "A Contextualist Theory of Epistemic Justification." *American Philosophical Quarterly* 15:213–19.

Ball, Brian and Blome-Tillmann, Michael. 2014. "Counter Closure and Knowledge Despite Falsehood." *Philosophical Quarterly* 64:552–568.

Baumann, Peter. 2008. "Contextualism and the Factivity Problem." *Philosophy and Phenomenological Research* 76:580–602.

—. 2010. "Factivity and Contextualism." *Analysis* 70:82–89.

Bennett, Jonathan. 2003. *A Philosophical Guide to Conditionals*. Oxford University Press.

Benton, Matthew A. 2014. "Gricean Quality." *Noûs* Online Early.

Bird, Alexander. 2005. "Abductive Knowledge and Holmesian Inference." In Tamar Szabo Gendler and John Hawthorne (eds.), *Oxford Studies in Epistemology*, volume 1, 1–31. Oxford University Press.

—. 2007. "Justified Judging." *Philosophy and Phenomenological Research* 74:81–110.

Blackson, Thomas A. 2004. "An Invalid Argument for Contextualism." *Philosophy and Phenomenological Research* 68:344–345.

Blome-Tillmann, Michael. 2008. "Contextualism, Safety and Epistemic Relevance." *Philosophical studies* 143:383–394.

—. 2009. "Knowledge and Presuppositions." *Mind* 118:241–294.

—. 2014. *Knowledge and Presuppositions*. Oxford University Press.

—. 2015a. "Ignorance, Presuppositions, and the Simple View." *Mind* Online:1–10.

—. 2015b. "Sensitivity Actually." *Philosophy and Phenomenological Research* Early View.

BonJour, Laurence. 2010. *Epistemology: Classic Problems and Contemporary Responses*. Rowman & Littlefield Publishers, Inc., 2nd edition.

BonJour, Laurence and Sosa, Ernest. 2003. *Epistemic Justification: Internalism Vs. Externalism, Foundations Vs. Virtues*. Blackwell Pub.

Borg, Emma. 2004. *Minimal Semantics*. Oxford University Press.

Boyd, Kenneth. 2015. "Pragmatic Encroachment and Epistemically Responsible Action." *Synthese* Online.

Braun, David. 2015. "Indexicals." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. `http://plato.stanford.edu/entries/indexicals/`, spring 2015 edition.

Brendel, Elke. 2005. "Why Contextualists Cannot Know They Are Right: Self-Refuting Implications of Contextualism." *Acta Analytica* 20:38–55.

Brogaard, Berit. 2008. "In Defence of a Perspectival Semantics for 'Know'." *Australasian Journal of Philosophy* 86:439–459.

—. 2012. *Transient Truths: An Essay in the Metaphysics of Propositions*. OUP USA.

Brogaard, Berit and Salerno, Joe. 2008. "Counterfactuals and Context." *Analysis* 68:39–46.

Bronfman, Aaron and Dowell, Janice J. L. 2016. "Contextualism About Deontic Conditionals." In Nate Charlow and Matthew Crisman (eds.), *Deontic Modality*. Oxford University Press.

Brown, Jessica. 2005. "Doubt, Circularity and the Moorean Response to the Sceptic." *Philosophical Perspectives* 19:1–14.

—. 2006. "Contextualism and Warranted Assertibility Manoeuvres." *Philosophical Studies* 130:407–435.

—. 2008a. "Knowledge and Practical Reason." *Philosophy Compass* 3:1135–1152.

—. 2008b. "The Knowledge Norm for Assertion." *Philosophical Issues* 18:89–103.

—. 2008c. "Subject-Sensitive Invariantism and the Knowledge Norm for Practical Reasoning." *Noûs* 42:167–189.

—. 2012a. "Assertion and Practical Reasoning: Common or Divergent Epistemic Standards?" *Philosophy and Phenomenological Research* 84:123–157.

—. 2012b. "Practial Reasoning, Decision Theory and Anti-Intellectualism." *Episteme* 9:1–20.

—. 2016. "The Knowledge Norm of Belief, Justification and Excuse." Invited Symposium, APA Central Division Meeting.

Brueckner, Anthony. 1994. "The Shifting Content of Knowledge Attributions." *Philosophy and Phenomenological Research* 54:123–126.

Brueckner, Anthony and Buford, Christopher T. 2009. "Contextualism, SSI and the Factivity Problem." *Analysis* 69:431–438.

Buford, Christopher. 2009. "Contextualism, Closure, and the Knowledge Account of Assertion." *Journal of Philosophical Research* 34:111–121.

Burch-Brown, Joanna M. 2014. "Clues for Consequentialists." *Utilitas* 26:105–119.

Burge, Tyler. 1979. "Individualism and the Mental." *Midwest Studies in Philosophy* 4:73–122.

—. 1991. "Vision and Intentional Content." In Ernest LePore and Robert Van Gulick (eds.), *John Searle and His Critics*, 195–214. Blackwell.

—. 2013. *Cognition Through Understanding: Self-Knowledge, Interlocution, Reasoning, Reflection: Philosophical Essays, Volume 3*. OUP Oxford.

Byrne, Alex and Logue, Heather. 2009. *Disjunctivism: Contemporary Readings*. MIT Press.

Calrlson, Lauri. 1983. *Dialogue Games: An Approach to Discourse Analysis*. Reidel.

Cappelen, Herman. 2012. *Philosophy Without Intuitions*. Oxford University Press.

Cappelen, Herman and Hawthorne, John. 2009. *Relativism and Monadic Truth*. Oxford University Press.

Cappelen, Herman and Lepore, Ernie. 2005. *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*. Oxford: Blackwell.

Carr, Jennifer. 2015. "Subjective Ought." *Ergo, an Open Access Journal of Philosophy* 2.

Carroll, Lewis. 1895. "What the Tortoise Said to Achilles." *Mind* 4:278–280.

Carter, J. Adam. 2013. "Disagreement, Relativism and Doxastic Revision." *Erkenntnis* 1–18.

Carter, J. Adam and Nickel, Philip J. 2014. "On Testimony and Transmission." *Episteme* 11:145–155.

Chalmers, David J. 2004. "The Representational Character of Experience." In Brian Leiter (ed.), *The Future for Philosophy*, 153–181. Oxford University Press.

Chan, Timothy. 2013. *The Aim of Belief*. Oxford University Press.

Chisholm, Roderick. 1942. "The Problem of the Speckled Hen." *Mind* 51:368–373.

Churchland, Patricia S. 1986. *Neurophilosophy: Toward A Unified Science of the Mind-Brain*. MIT Press.

Churchland, Paul M. 1981. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78:67–90.

Clarke, Roger. 2013. "Belief Is Credence One (In Context)." *Philosophers' Imprint* 13:1–18.

—. 2015. "Preface Writers Are Consistent." *Pacific Philosophical Quarterly* Early View.

—. 2016. "Assertion, Belief and Context." Online (2 June, 2016) from `http://www.rogerclarke.org/research`.

—. 2017. "Contextualism about Belief Ascriptions." In Ichikawa (2017).

Cohen, Stewart. 1984. "Justification and Truth." *Philosophical Studies* 46:279–95.

—. 1998. "Contextualist Solutions to Epistemological Problems: Scepticism, Gettier, and the Lottery." *Australasian Journal of Philosophy* 76:289–306.

—. 1999. "Contextualism, Skepticism, and the Structure of Reasons." *Philosophical Perspectives* 13:57–89.

—. 2005. "Knowledge, Speaker and Subject." *Philosophical Quarterly* 55:199–212.

—. 2016. "Theorizing about the Epistemic." *Inquiry* In Press. Draft online (2 June, 2016) from `http://www.stew-cohen.com/papers/`.

Comesaña, Juan and Kantin, Holly. 2010. "Is Evidence Knowledge?" *Philosophy and Phenomenological Research* 80:447–454.

Conee, E. and Feldman, R. 1998. "The Generality Problem for Reliabilism." *Philosophical Studies* 89:1–29.

Craig, Edward. 1990. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford University Press.

Davidson, Donald. 1963. "Actions, Reasons, and Causes." *Journal of Philosophy* 60:685–700.

Davis, Wayne A. 2007. "Knowledge Claims and Context: Loose Use." *Philosophical Studies* 132:395–438.

—. 2017. "Loose Use and Belief Variation." In Ichikawa (2017).

Dennett, Daniel C. 1987. *The Intentional Stance*. MIT Press.

DeRose, Keith. 1992. "Contextualism and Knowledge Attributions." *Philosophy and Phenomenological Research* 52:913–29.

—. 1995. "Solving the Skeptical Problem." *The Philosophical Review* .

—. 1999. "Can It Be That It Would Have Been Even Though It Might Not Have Been?" *Noûs* 33:385–413.

—. 2002. "Assertion, Knowledge, and Context." *The Philosophical Review* 111:167–203.

—. 2005. "The Ordinary Language Basis for Contextualism, and the New Invariantism." *Philosophical Quarterly* 55:172–198.

—. 2009. *The Case for Contextualism*. Oxford University Press.

Dimmock, Paul and Huvenes, Torfinn Thomesen. 2014. "Knowledge, Conservatism, and Pragmatics." *Synthese* 191:3239–3269.

Dodd, Dylan. 2007. "Why Williamson Should Be a Sceptic." *Philosophical Quarterly* 57:635–649.

Douven, Igor. 2008. "Knowledge and Practical Reasoning." *Dialectica* 62:101–118.

—. 2012. "The Lottery Paradox and the Pragmatics of Belief." *Dialectica* 66:351–373.

Dowell, Janice J. L. 2011. "A Flexible Contextualist Account of Epistemic Modals." *Philosophers' Imprint* 11:1–25.

—. 2012. "Contextualist Solutions to Three Puzzles About Practical Conditionals." In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, volume 7*. Oxford.

—. 2017. "Contextualism About Epistemic Modals." In Ichikawa (2017).

Dretske, Fred I. 1970. "Epistemic Operators." *Journal of Philosophy* 67:1007–1023.

—. 1981. *Knowledge and the Flow of Information*. MIT Press.

Driver, Julia. 2001. *Uneasy Virtue*. Cambridge University Press.

Dummett, Michael. 1973. *Frege: Philosophy of Language*. Duckworth.

Dutant, Julien. 2015. "The Legend of the Justified True Belief Analysis." *Philosophical Perspectives* 29:95–145.

Edgington, Dorothy. 2011. "Causation First: Why Causation is Prior to Counterfactuals." In Christoph Hoerl, Teresa McCormack, and Sarah R. Beck (eds.), *Understanding Counterfactuals, Understanding Causation: Issues in Philosophy and Psychology*. Oxford University Press.

Egan, Andy, Hawthorne, John, and Weatherson, Brian. 2005. "Epistemic Modals in Context." In G. Preyer and G. Peter (eds.), *Contextualism in Philosophy*, 131–170. Oxford University Press.

Engel, Mylan. 1992. "Personal and Doxastic Justification in Epistemology." *Philosophical Studies* 67:133–150.

Enoch, David and Schechter, Joshua. 2008. "How Are Basic Belief-Forming Methods Justified?" *Philosophy and Phenomenological Research* 76:547–579.

Fantl, Jeremy and McGrath, Matthew. 2009. *Knowledge in an Uncertain World*. Oxford University Press.

Fantl, Jeremy and Mcgrath, Matthew. 2012. "Contextualism and Subject-Sensitivity." *Philosophy and Phenomenological Research* 84:693–702.

Faulkner, Paul. 2006. "Understanding Knowledge Transmission." *Ratio* 19:156–175.

Feldman, Richard. 2005. "Justification is Internal." In Steup Matthias and Sosa Ernest (eds.), *Contemporary Debates in Epistemology*, 270–84. Blackwell.

Feldman, Richard and Conee, Earl. 1985. "Evidentialism." *Philosophical Studies* 48:15–34.

—. 2001. "Internalism Defended." *American Philosophical Quarterly* 38:1–18.

Fine, Cordelia. 2006. "Is the Emotional Dog Wagging its Rational Tail, or Chasing It?" *Philosophical Explorations* 9:83–98.

Fine, Kit. 1975. "Critical Notice of Lewis, Counterfactuals." *Mind* 84:451–458.

Fitelson, Branden. 2010. "Strengthening the Case for Knowledge From Falsehood." *Analysis* 70:666–669.

Fogal, Daniel. 2016. "Reasons, Reason, and Context." In Errol Lord and Barry Maguire (eds.), *Weighing Reasons*, 74–103. Oxford University Press.

Franklin, Benjamin. 1956. *Benjamin Franklin: A Selection from His Personal Letters*. Yale University Press.

Freitag, Wolfgang. 2011. "Epistemic Contextualism and the Knowability Problem." *Acta Analytica* 26:273–284.

Fricker, Elizabeth. 2012. "Stating and Insinuating." *Aristotelian Society Supplementary Volume* 86:61–94.

Fricker, Miranda. 2009. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.

Fumerton, Richard. 1995. *Metaepistemology and Skepticism*. Rowman & Littlefield.

Gauker, Christopher. 2010. "Global Domains Versus Hidden Indexicals." *Journal of Semantics* 27:243–270.

Gerken, Mikkel. 2011. "Warrant and Action." *Synthese* 178:529–547.

—. 2013. "Epistemic Focal Bias." *Australasian Journal of Philosophy* 91:41–61.

Gettier, Edmund L. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23:121–123.

Gibbons, John. 2013. *The Norm of Belief*. Oxford: Oxford University Press.

Ginet, Carl. 1975. *Knowledge, Perception, and Memory*. D. Reidel Pub. Co.

Ginzburg, Jonathan. 1996. "The Semantics of Interrogatives." In Shalom Lappen (ed.), *The Handbook of Contemporary Semantic Thoery*, 385–422. Basil Blackwell.

Goldberg, Sanford C. 2015. *Assertion: On the Philosophical Significance of Assertoric Speech*. Oxford University Press.

Goldman, Alvin and McGrath, Matthew. 2014. *Epistemology: An Introduction*. Oxford.

Goldman, Alvin I. 1979. "What is Justified Belief?" In G. S. Pappas (ed.), *Justification and Knowledge*, 1–25. Boston: D. Reidel.

—. 1986. *Epistemology and Cognition*. Harvard University Press.

—. 1988. "Strong and Weak Justification." *Philosophical Perspectives* 2:51–69.

—. 1999. "Internalism Exposed." *Journal of Philosophy* 96:271–293.

—. 2007. "Philosophical Intuitions: Their target, Their Source, and Their Epistemic Status." *Grazer Philosophische Studien* 1–26.

—. 2009. "Williamson on Knowledge and Evidence." In Patrick Greenough and Duncan Pritchard (eds.), *Williamson on Knowledge*, 73–91. Oxford University Press.

—. 2011. "Reliabilism." In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. `http://plato.stanford.edu/entries/reliabilism/`, spring 2011 edition.

Grimm, Stephen R. 2011. "On Intellectualism in Epistemology." *Mind* 120:705–733.

Haddock, Adrian and Macpherson, Fiona. 2008. *Disjunctivism: Perception, Action, Knowledge*. Oxford University Press.

Haidt, Jonathan. 2001. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108:814–34.

Halliday, Daniel. 2007. "Contextualism, Comparatives and Gradability." *Philosophical Studies* 132:381–393.

Harman, Gilbert. 1986. *Change in View: Principles of Reasoning*. The MIT Press.

Haslanger, Sally. 2000. "Gender and Race: (What) Are They? (What) Do We Want Them to Be?" *Noûs* 34:31–55.

—. 2006. "What Good are Our Intuitions?" *Aristotelian Society Supplementary Volume* 80:89–118.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford University Press.

—. 2012. "Some Comments on Fricker's 'Stating and Insinuating'." *Aristotelian Society Supplementary Volume* 86:95–108.

Hawthorne, John and Magidor, Ofra. 2016. "Reflections on Reasons." In Daniel Star (ed.), *The Oxford Handbook of Reasons and Normativity*. Oxford University Press. In Press. Draft online (2 June, 2016) from `http://users.ox.ac.uk/~ball1646/publications.html`.

Hawthorne, John and Stanley, Jason. 2008. "Knowledge and Action." *Journal of Philosophy* 105:571–590.

Heylen, Jan. 2016. "Being in a Position to Know and Closure." *Thought: A Journal of Philosophy* 5:63–67. ISSN 2161-2234. doi:10.1002/tht3.194.

Hill, Christopher S. and Schechter, Joshua. 2007. "Hawthorne's Lottery Puzzle and the Nature of Belief." *Philosophical Issues* 17:1020–122.

Holman, Emmett L. 2002. "Color Eliminativism and Color Experience." *Pacific Philosophical Quareterly* 83:38–56.

Howard-Snyder, Frances. 1997. "The Rejection of Objective Consequentialism." *Utilitas* 9:241–248.

Huemer, Michael. 1999. "The Problem of Memory Knowledge." *Pacific Philosophical Quarterly* 80:346–357.

Huvenes, Torfinn Thomesen. 2014. "Disagreement Without Error." *Erkenntnis* 79:143–154.

—. 2017. "On Disagreement." In Ichikawa (2017).

Hyman, John. 1999. "How Knowledge Works." *Philosophical Quarterly* 50:433–451.

Ichikawa, Jonathan Jenkins. 2009a. "Explaining Away Intuitions." *Studia Philosophica Estonica* 2:94–116.

—. 2009b. "Review of Keith DeRose, *The Case for Contextualism: Knowledge, Skepticism, and Context, Vol. 1*." *Notre Dame Philosophical Reviews* 2009.

—. 2011a. "Quantifiers and Epistemic Contextualism." *Philosophical Studies* 155:383–398.

—. 2011b. "Quantifiers, Knowledge, and Counterfactuals." *Philosophy and Phenomenological Research* 82:287–313.

—. 2012. "Knowledge Norms and Acting Well." *Thought: A Journal of Philosophy* .

—. 2013. "Basic Knowledge and Contextualist "E = K"." *Thought: A Journal of Philosophy* 2:282–292.

—. 2014. "Justification is Potential Knowledge." *Canadian Journal of Philosophy* 44:184–206.

—. 2015. "Ignorance and Presuppositions." *Mind* Online.

—. 2016a. "Basic Knowledge First." *Episteme* In Press. Draft online (1 July, 2016 from `http://jichikawa.net/research/`).

—. 2016b. "Modals and Modal Epistemology." In Amy Kind and Peter Kung (eds.), *Knowledge Through Imagination*, 124–44. Oxford University Press.

Ichikawa, Jonathan Jenkins (ed.). 2017. *Routledge Handbook of Epistemic Contextualism*. Taylor & Francis.

Ichikawa, Jonathan Jenkins, Jarvis, Benjamin, and Rubin, Katherine. 2012. "Pragmatic Encroachment and Belief-Desire Psychology." *Analytic Philosophy* 53:327–343.

Ichikawa, Jonathan Jenkins and Jarvis, Benjamin W. 2009. "Thought-Experiment Intuitions and Truth in Fiction." *Philosophical Studies* 142:221–46.

—. 2013. *The Rules of Thought*. Oxford: Oxford University Press.

—. 2016. "Hybrid Virtue Epistemology and the A Priori." In Dylan Dodd and Elia Zardini (eds.), *The A Priori: Its Significance, Grounds, and Extent*. Oxford University Press. In Press.

Ichikawa, Jonathan Jenkins and Jenkins, C. S. I. 2016. "On Putting Knowledge 'First'." In J. Adam Carter, Emma Gordon, and Benjamin W. Jarvis (eds.), *Knowledge First*. Oxford University Press. In Press. Draft online (2 June, 2016) from `http://jichikawa.net/research/`.

Jackson, Frank. 1991. "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101:461–482.

Jäger, Christoph. 2012. "Contextualism and the Knowledge Norm of Assertion." *Analysis* 72:491–498.

Jenkins, C. S. 2006. "Knowledge and Explanation." *Canadian Journal of Philosophy* 36:137–163.

Jenkins, C. S. I. 2008. "Romeo, René, and the Reasons Why: What Explanation Is." *Proceedings of the Aristotelian Society* 108:61–84.

—. 2013. "Explanation and Fundamentality." In *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-Dependence (Basic Philosophical Concepts)*, 211–42. Philosophia Verlag.

Jenkins, C. S. I. and Nolan, Daniel. 2010. "Maximising, Satisficing and Context." *Noûs* 44:451–468.

Johnson, Casey Rebecca. 2015. "Testimony and the Constitutive Norm of Assertion." *International Journal of Philosophical Studies* 23:356–375.

Joyce, Jim. 2004. "Williamson on Evidence and Knowledge." *Philosophical Books* 45:296–305.

Joyce, Richard. 2015. "Moral Anti-Realism." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. `http://plato.stanford.edu/entries/moral-anti-realism/`, fall 2015 edition.

Kalkanis, George, Hadzidaki, Pandora, and Stavrou, Dimitrios. 2003. "An Instructional Model for a Radical Conceptual Change towards Quantum Mechanics Concepts." *Science Education* 87:257–280. ISSN 1098-237X. doi:10.1002/sce.10033.

Kaplan, David. 1977. "Demonstratives." In *Themes From Kaplan*, 481–563. Oxford University Press.

Kaplan, Mark. 1985. "It's Not What You Know That Counts." *Journal of Philosophy* 82:350–363.

Kearns, Stephen and Star, Daniel. 2009. "Reasons as Evidence." *Oxford Studies in Metaethics* 4:215–42.

Kelp, Christoph. 2011. "Not Without Justification." *Dialectica* 65:581–595.

—. 2014. "No Justification for Lottery Losers." *Pacific Philosophical Quarterly* 95:205–217.

—. 2016. "Justified Belief: Knowledge First-Style." *Philosophy and Phenomenological Research* Online Early:1–22.

Khoo, Justin. 2017. "The Disagreement Challenge to Contextualism." In Ichikawa (2017).

King, Jeffrey. 2005. "Semantics, Pragmatics, and the Role of Semantic Content." In Zoltan Szabo (ed.), *Semantics Versus Pragmatics*, 111–164. Oxford University Press.

Kneale, William Calvert. 1949. *Probability and Induction*. Clarendon Press.

Kolodny, Niko and MacFarlane, John. 2010. "Ifs and Oughts." *Journal of Philosophy* 107:115–143.

Kompa, Nikola. 2002. "The Context Sensitivity of Knowledge Ascriptions." *Grazer Philosophische Studien* 64:1–18.

Kornblith, Hilary. 2000. "The Contextualist Evasion of Epistemology." *Noûs* 34:24–32.

—. 2004. "Does Reliabilism Make Knowledge Merely Conditional?" *Philosophical Issues* 14:185–200.

Kratzer, Angelika. 2012. "Conditionals." In *Modals and Conditionals*, 86–108. Oxford University Press.

Kripke, Saul A. 1980. *Naming and Necessity*. Harvard University Press.

—. 1982. *Wittgenstein on Rules and Private Language*. Harvard University Press.

—. 2011. "Nozick on Knowledge." In Saul A. Kripke (ed.), *Philosophical Troubles. Collected Papers Vol I*, 162–224. Oxford University Press.

Kyburg, Henry Ely. 1961. *Probability and the Logic of Rational Belief*. Middletown, Conn.,Wesleyan University Press.

L'Abate, Luciano and Bryson, Charles H. 1994. *A Theory of Personality Development*. Wiley-Blackwell.

Lackey, Jennifer. 2007. "Norms of Assertion." *Noûs* 41:594–626.

—. 2008. *Learning From Words: Testimony as a Source of Knowledge*. Oxford University Press.

—. 2010. "Acting on Knowledge." In Lackey (2014b), 361–382.

—. 2014a. "A Deflationary Account of Group Testimony." In Lackey (2014b), 64–94.

—. 2014b. *Essays in Collective Epistemology*. Oxford University Press.

Lasonen-Aarnio, Maria. 2008. "Single Premise Deduction and Risk." *Philosophical Studies* 141:157–73.

Lenman, James. 2000. "Consequentialism and Cluelessness." *Philosophy and Public Affairs* 29:342–370.

Lewis, David K. 1973. *Counterfactuals*. Oxford: Blackwell Publishers.

—. 1979a. "Counterfactual Dependence and Time's Arrow." *Noûs* 13:455–476.

—. 1979b. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8:339–59.

—. 1984. "Putnam's Paradox." *Australasian Journal of Philosophy* 62:221–236.

—. 1986. *On the Plurality of Worlds*. Wiley-Blackwell.

—. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy* 74:549–567.

Lewis, Karen S. 2015. "Elusive Counterfactuals." *Noûs* 50:286–313.

—. 2016. "Counterfactual Discourse in Context." Retrieved 31 May, 2016 from `http://www.columbia.edu/~kl2663/`.

—. 2017. "Counterfactuals and Knowledge." In Ichikawa (2017).

Littlejohn, Clayton. 2012. *Justification and the Truth-Connection*. Cambridge University Press.

—. 2013. "No Evidence is False." *Acta Analytica* 28:145–159.

—. 2014. "The Unity of Reason." In Clayton Littlejohn and John Turri (eds.), *Epistemic Norms: New Essays on Action, Belief and Assertion*, 135–154. Epistemic Norms: New Essays on Action, Belief and Assertion.

Locke, Dustin. 2014. "Knowledge Norms and Assessing Them Well." *Thought: A Journal of Philosophy* 3:80–89.

—. 2015. "Practical Certainty." *Philosophy and Phenomenological Research* 90:72–95.

Locke, John. 2008. *An Essay Concerning Human Understanding*. Oxford University Press.

Luzzi, Federico. 2010. "Counter-Closure." *Australasian Journal of Philosophy* 88:673–683.

—. 2011. "Contextualism and Counter-Closure." *Dialectica* 66:187–199.

MacFarlane, John. 2005. "The Assessment Sensitivity of Knowledge Attributions." In Tamar Szabo Gendler John Hawthorne (ed.), *Oxford Studies in Epistemology*, volume 1, 197–234. Oxford University Press.

—. 2007. "Relativism and Disagreement." *Philosophical Studies* 132:17–31.

Macpherson, Fiona. 2010. "A Disjunctive Theory of Introspection: A Reflection on Zombies and Anton's Syndrome." *Philosophical Issues* 20:226–265.

Manne, Kate. 2014. "Internalism About Reasons: Sad but True?" *Philosophical Studies* 167:89–117.

Marion, Mathieu. 2000. "Oxford Realism: Knowledge and Perception II." *British Journal for the History of Philosophy* 8:485–519.

Markosian, Ned. 2014. "Do You Know that You are not a Brain in a Vat?" *Logos & Episteme* 5:161–181.

Maund, J. Barry. 1995. *Colours: Their Nature and Representation*. Cambridge University Press.

McDowell, John. 1982. "Criteria, Defeasibility, and Knowledge." *Proceedings of the British Academy* 68:455–79.

McGlynn, Aidan. 2012. "Justification as 'Would-Be' Knowledge." *Episteme* 9:361–76.

—. 2014. *Knowledge First?* Palgrave Macmillian.

McGrath, Matthew. 2010. "Contextualism and Intellectualism." *Philosophical Perspectives* 24:383–405.

McKenna, Robin. 2014. "Shifting Targets and Disagreements." *Australasian Journal of Philosophy* 92:725–742.

McKinnon, Rachel. 2015. *The Norms of Assertion: Truth, Lies, and Warrant*. Palgrave Macmillan.

Mehtsun, Winta T., Ibrahim, Andrew M., Diener-West, Marie, Pronovost, Peter J., and Makary, Martin A. 2013. "Surgical Never Events in the United States." *Surgery: Official Journal for the Society of University Surgeons* 153.

Miracchi, Lisa. 2015a. "Competence to Know." *Philosophical Studies* 172:29–56.

—. 2015b. "Knowledge Is All You Need." *Philosophical Issues* 25:353–378.

—. 2016. "Competent Perspectives and the New Evil Demon Problem." In *The New Evil Demon: New Essays on Knowledge, Justification and Rationality*. Oxford University Press. In Press. Draft online (2 June, 2016 from `http://miracchi.wix.com/lisamiracchi`.

Moon, Andrew. 2012. "Three Forms of Internalism and the New Evil Demon Problem." *Episteme* 9:345–360.

Moore, George Edward. 1939. "Proof of an External World." *Proceedings of the British Academy* 25:273–300.

—. 1959. "Four Forms of Scepticism." In *Philosophical Papers*. Georg Allen & Unwin.

—. 1962/2013. *Commonplace Book: 1919–1953*. Taylor & Francis.

Moss, Sarah. 2012. "On the Pragmatics of Counterfactuals." *Noûs* 46:561–86.

Mulligan, Kevin. 2014. "Knowledge First: A German Folly?" In J. Dutant, D. Fassio, and A. Meylan (eds.), *Liber Amicorum Pascal Engel*, 380–400. University of Geneva.

Myers-Schulz, Blake and Schwitzgebel, Eric. 2013. "Knowing That P Without Believing That P." *Noûs* 47:371–384.

Nagel, Jennifer. 2010. "Epistemic Anxiety and Adaptive Invariantism." *Philosophical Perspectives* 24:407–35.

—. 2011. "The Psychological Basis of the Harman-Vogel Paradox." *Philosophers' Imprint* 11:1–28.

—. 2013. "Knowledge as a Mental State." *Oxford Studies in Epistemology* 4:275–310.

—. 2014. *Knowledge: A Very Short Introduction*. Oxford University Press.

Nelkin, Dana K. 2000. "The Lottery Paradox, Knowledge, and Rationality." *Philosophical Review* 109:373–409.

Neta, Ram. 2003. "Contextualism and the Problem of the External World." *Philosophy and Phenomenological Research* 66:1–31.

—. 2008a. "In Defense of Disjunctivism." In Fiona Macpherson and Adrian Haddock (eds.), *Disjunctivism: Perception, Action, Knowledge*, 311–29. Oxford University Press.

—. 2008b. "What Evidence Do You Have?" *British Journal for the Philosophy of Science* 59:89–119.

—. 2009. "Treating Something as a Reason for Action." *Noûs* 43:684–699.

Norcross, Alastair. 1990. "Consequentialism and the Unforeseeable Future." *Analysis* 50:253–256.

Nozick, Robert. 1981. *Philosophical Explanations*. Harvard University Press.

Oakley, I. T. 2001. "A Skeptic's Reply to Lewisian Contextualism." *Canadian Journal of Philosophy* 31:309–332.

Pagin, Peter. 2015. "Problems with Norms of Assertion." *Philosophy and Phenomenological Research* Early View.

Pappas, George. 2014. "Internalist vs. Externalist Conceptions of Epistemic Justification." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. `http://plato.stanford.edu/archives/fall2014/entries/justep-intext/`, fall 2014 edition.

Pappas, George S. 1980. "Lost Justification." *Midwest Studies in Philosophy* 5:127–134.

Partee, Barbara Hall. 1973. "Some Structural Analogies Between Tenses and Pronouns in English." *Journal of Philosophy* 70:601–609.

Pettit, P. 2003. "Groups with Minds of their Own." In Frederick F. Schmitt (ed.), *Socializing Metaphysics: The Nature of Social Reality*, 167–93. Rowman & Littlefield Publishers, Inc.

Portner, Paul. 2009. *Modality*. Oxford University Press.

Pritchard, Duncan. 2002. "Two Forms of Epistemological Contextualism." *Grazer Philosophische Studien* 64:19–55.

—. 2005. "Neo-Mooreanism Versus Contextualism." *Grazer Philosophische Studien* 69:1–24.

—. 2007. "How to Be a Neo-Moorean." In Sanford Goldberg (ed.), *Internalism and Externalism in Semantics and Epistemology*, 68–99. Oxford University Press.

—. 2012. *Epistemological Disjunctivism*. Oxford University Press.

Radford, Colin. 1966. "Knowledge—By Examples." *Analysis* 27:1–11.

Reed, Baron. 2010. "A Defense of Stable Invariantism." *Noûs* 44:224–244.

Reynolds, Steven L. 1998. "Evaluational Illusions and Skeptical Arguments." *Philosophy and Phenomenological Research* 58:529–558.

—. 2013. "Justification as the Appearance of Knowledge." *Philosophical Studies* 163:367–383.

Richard, Mark. 2004. "Contextualism and Relativism." *Philosophical Studies* 119:215–242.

Rizzieri, Aaron. 2009. "Evidence does not equal knowledge." *Philosophical studies* 153:235–242.

Roberts, Craige. 2004. "Context in Dynamic Interpretation." In Laurence R. Horn and Gregory Ward (eds.), *The Handbook of Pragmatics*, 197–220. Basil Blackwell.

Rogers, M.L., Cook, R.I., Bower, R., Molloy, M., and Render, M.L. 2004. "Barriers to Implementing Wrong Site Surgery Guidelines: A Cognitive Work Analysis." *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 34:757–763. ISSN 1083-4427. doi:10.1109/TSMCA.2004.836805.

Roland, Jeffrey and Cogburn, Jon. 2011. "Anti-Luck Epistemologies and Necessary Truths." *Philosophia* 39:547–561.

Roush, Sherrilyn. 2005. *Tracking Truth: Knowledge, Evidence, and Science*. Oxford University Press.

Rubin, Katherine. 2015. "Total Pragmatic Encroachment and Epistemic Permissiveness." *Pacific Philosophical Quarterly* 96:12–38.

Russell, Bertrand. 2009 [1948]. *Human Knowledge: Its Scope and Limits*. Routledge.

Russell, Gillian and Doris, John M. 2008. "Knowledge by Indifference." *Australasian Journal of Philosophy* 86:429–37.

Schaffer, Jonathan. 2004. "From Contextualism to Contrastivism." *Philosophical Studies* 119:73–109.

—. 2005. "What Shifts? : Thresholds, Standards, or Alternatives?" In Gerhard Preyer and Georg Peter (eds.), *Contextualism in Philosophy: Knowledge, Meaning, and Truth*. Oxford University Press.

—. 2006. "The Irrelevance of the Subject: Against Subject-Sensitive Invariantism." *Philosophical Studies* 127:87–107.

—. 2008. "Knowledge in the Image of Assertion." *Philosophical Issues* 18:1–19.

—. 2012. "Necessitarian Propositions." *Synthese* 189:119–162.

—. 2015. "Lewis on Knowledge Ascriptions." In Barry Loewer and Jonathan Schaffer (eds.), *A Companion to David Lewis*, chapter 30, 473–490. John Wiley & Sons, Inc.

Schiffer, Stephen. 2007. "Interest-Relative Invariantism." *Philosophy and Phenomenological Research* 75:188–195.

Schroeder, Mark. 2008. "Having Reasons." *Philosophical Studies* 139:57–71.

—. 2011. "What Does it Take to 'Have' a Reason?" In Andrew Reisner and Asbjørn Steglich-Peterson (eds.), *Reasons for Belief*, 201–22. Cambridge University Press.

Schwitzgebel, Eric. 2011. *Perplexities of Consciousness*. The MIT Press.

Sgaravatti, Daniele. 2013. "In Conversation with the Skeptic: Contextualism and the Raising of Standards." *International Journal for the Study of Skepticism* 3:97–118.

Shapiro, Stewart. 2014. *Varieties of Logic*. Oxford University Press.

Shoemaker, Sydney. 1990. "Qualities and Qualia: What's in the Mind?" *Philosophy and Phenomenological Research* 50:109–131.

—. 1994. "Lecture III: The Phenomenal Character of Experience – Self Knowledge and Inner Sense." *Philosophy and Phenomenological Research* 54:291–314.

—. 2006. "On the Way Things Appear." In Tamar S. Gendler and John Hawthorne (eds.), *Perceptual Experience*, 461–480. Oxford University Press.

Sider, Theodore. 2011. *Writing the Book of the World*. Oxford University Press.

Siegal, Susanna. 2015. "The Contents of Perception." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2015 Edition)*. `http://plato.stanford.edu/archives/spr2015/entries/perception-contents/`, spring 2015 edition.

Silins, Nicholas. 2011. "Seeing Through the 'Veil of Perception'." *Mind* 120:329–367.

Simons, Mandy. 2007. "Observations on Embedding Verbs, Evidentiality, and Presupposition." *Lingua* 117:1034–56.

Smith, Martin. 2010. "What Else Justification Could Be." *Noûs* 44:10–31.

Sosa, Ernest. 1999. "How to Defeat Opposition to Moore." *Philosophical Perspectives* 13:141–53.

—. 2000. "Skepticism and Contextualism." *Philosophical Issues* 10:1–18.

—. 2003. "Privileged Access." In Quentin Smith and Aleksandar Jokic (eds.), *Consciousness: New Philosophical Perspectives*, 238–251. Oxford University Press.

—. 2009. "A Defense of the Use of Intuitions in Philosophy." In Dominic Murphy and Michael Bishop (eds.), *Stich and His Critics*, 101–12. Wiley-Blackwell.

Stalnaker, Robert C. 1974. "Pragmatic Presuppositions." In Robert Stalnaker (ed.), *Context and Content*, 47–62. Oxford University Press.

—. 1981a. "A Defense of Conditional Excluded Middle." In William Harper, Robert C. Stalnaker, and Glenn Pearce (eds.), *Ifs*, 87–104. Reidel.

—. 1981b. *Inquiry*. Cambridge, MA: The MIT Press.

—. 1998. "On the Representation of Context." *Journal of Logic, Language and Information* 7:3–19.

—. 2002. "Common Ground." *Linguistics and Philosophy* 25:701–721.

Stanley, Jason. 2005. *Knowledge and Practical Interests*. Oxford University Press.

—. 2008. "Knowledge and Certainty." *Philosophical Issues* 18:35–57.

Stanley, Jason and Szabó, Zoltán Gendler. 2000. "On Quantifier Domain Restriction." *Mind and Language* 15:219–61.

Star, Daniel. 2008. "Moral Knowledge, Epistemic Externalism, and Intuitionism." *Ratio* 21:329–343.

—. 2015. *Knowing Better*. Oxford University Press.

Steup, Matthias. 2001. *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford University Press.

Stone, Jim. 2007. "Contextualism and Warranted Assertion." *Pacific Philosophical Quarterly* 88:92–113.

Sutton, Jonathan. 2007. *Without Justification*. The MIT Press.

Thau, Michael. 2002. *Consciousness and Cognition*. Oxford University Press.

Turri, John. 2014a. "Skeptical Appeal: The Source-Content Bias." *Cognitive Science* 38:307–324.

—. 2014b. "You Gotta Believe." In Clayton Littlejohn and John Turri (eds.), *Epistemic Norms: New Essays on Action, Belief and Assertion*, chapter 10, 193–9. Oxford University Press.

—. 2015. "Knowledge and the Norm of Assertion: A Simple Test." *Synthese* 192:385–392.

Unger, Peter K. 1975. *Ignorance: A Case for Scepticism*. 105. Oxford University Press.

—. 1995. "Contextual Analysis in Ethics." *Philosophy and Phenomenological Research* 55:1–26.

van Eijck, Jan. 2010. "Dynamic Semantics." In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. `http://plato.stanford.edu/entries/dynamic-semantics/`, winter 2012 edition.

von Fintel, Kai. 2001. "Counterfactuals in a Dynamic Context." In *Ken Hale: A Life in Language*, chapter 3, 123–152. The MIT Press.

von Fintel, Kai and Gillies, Anthony S. 2007. "An Opinionated Guide to Epistemic Modality." In Tamar Gendler and John Hawthorne (eds.), *Oxford Studies in Epistemology 2*, volume 2, 32–62. Oxford.

Warfield, Ted A. 2005. "Knowledge From Falsehood." *Philosophical Perspectives* 19:405–8211.

Weatherson, Brian. 2003. "What Good are Counterexamples?" *Philosophical Studies* 115:1–31.

—. 2005. "Can We Do Without Pragmatic Encroachment?" *Philosophical Perspectives* 19:417–443.

—. 2009. "E ≠ K." Online (2 June, 2016) from `http://brian.weatherson.org/EK.pdf`.

—. 2011. "Defending Interest-Relative Invariantism." *Logos & Episteme* 2:591–609.

—. 2012. "Knowledge, Bets, and Interests." In Jessica Brown and Mikkel Gerken (eds.), *Knowledge Ascriptions*, 75–103. Oxford University Press.

Wedgwood, Ralph. 2002. "Internalism Explained." *Philosophy and Phenomenological Research* 65:349–369.

—. 2008. "Contextualism about Justified Belief." *Philosophers' Imprint* 8.

Weisberg, Jonathan. 2013. "Knowledge in Action." *Philosophers' Imprint* 13.

Whiting, Daniel. 2013. "Nothing but the Truth: On the Norms and Aims of Belief." In Timothy Chan (ed.), *The Aim of Belief*. Oxford University Press.

Wierzbicka, Anna. 1996. *Semantics: Primes and Universals*. Oxford University Press.

Williams, Bernard. 1973. "Deciding to Believe." In Bernard Williams (ed.), *Problems of the Self*, 136–51. Cambridge University Press.

—. 1979. "Internal and External Reasons." In Ross Harrison (ed.), *Rational Action*, 101–113. Cambridge University Press.

Williams, J. Robert G. 2010. "Defending Conditional Excluded Middle." *Noûs* 44:650–668.

Williams, Michael. 1991. *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*. B. Blackwell.

—. 2001a. "Contextualism, Externalism and Epistemic Standards." *Philosophical Studies* 103:1–23.

—. 2001b. *Problems of Knowledge: A Critical Introduction to Epistemology*. OUP Oxford.

—. 2004. "Scepticism and the Context of Philosophy." *Philosophical Issues* 14:456–475.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press. ISBN 0198250436 (alk. paper).

—. 2005a. "Contextualism, Subject-Sensitive Invariantism and Knowledge of Knowledge." *Philosophical Quarterly* 55:213–35.

—. 2005b. "Knowledge, Context, and the Agent's Point of View." In Gerhard Preyer and Georg Peter (eds.), *Contextualism in Philosophy*. Oxford University Press.

—. 2007. *The Philosophy of Philosophy*. Blackwell.

—. 2016. "Justifications, Excuses, and Sceptical Scenarios." In Julien Dutant and Fabian Dorsch (eds.), *The New Evil Demon*. Oxford University Press. In Press. Draft online (2 June, 2016) from `http://www.philosophy.ox.ac.uk/members/philosophy_panel/tim_williamson`.

Worsnip, Alex. 2015. "Belief, Credence, and the Preface Paradox." *Australasian Journal of Philosophy* Online Early:1–14.

Wright, Crispin J. G. 2000. "Cogency and Question-Begging: Some Reflections on McKinsey's Paradox and Putnam's Proof." *Philosophical Issues* 10:140–63.

—. 2002. "(Anti-)Sceptics Simple and Subtle: G. E. Moore and John McDowell." *Philosophy and Phenomenological Research* 65:330–348.

—. 2005. "Contextualism and Scepticism: Even-handedness, Factivity and Surreptitiously Raising Standards." *The Philosophical Quarterly* 55:236–62.

Wright, Sarah. 2010. "Virtues, Social Roles, and Contextualism." *Metaphilosophy* 41:95–114.

—. 2011. "Knowledge and Social Roles: A Virtue Approach." *Episteme* 8:99–111.

Wright, Wayne. 2003. "Projectivist Representationalism and Color." *Philosophical Psychology* 16:515–529.

Yourgrau, Palle. 1983. "Knowledge and Relevant Alternatives." *Synthese* 55:175–190.

Zagzebski, Linda Trinkaus. 1994. "The Inescapability of Gettier Problems." *The Philosophical Quarterly* 44:65–73.

—. 1996. *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press.