Belief in Context

by

Roger Clarke

B.Sc., The University of Toronto, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in
The Faculty of Graduate Studies
(Philosophy)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
December 2011
© Roger Clarke 2011
Abstract

I argue for a view I call sensitivism about belief. According to sensitivism, belief is sensitive to just those factors of context which epistemic contextualists claim are relevant to the semantics of words like “know”: in particular, whether an agent believes \( p \) depends on the not-\( p \) alternatives salient to the agent, and the practical importance of \( p \) for the agent. I argue for sensitivism about both outright belief and partial belief, and outline a sensitivist formal model of belief.

In chapter 1, I make a preliminary case for sensitivism, and for interest in sensitivism. After surveying some similar views in the literature, I present a scenario which is nicely explained by sensitivism, and which gives the view some intuitive plausibility. I also argue for the relevance of sensitivism to the debate over epistemic contextualism.

In chapter 2, I argue that we need sensitivism about outright belief if we want to maintain both a Stalnakerian picture of how assertion works, and the principle that an assertion that \( p \) is sincere if and only if the assertor believes that \( p \). I then outline a sensitivist formal model of outright belief. In chapter 3, I present a solution to the preface paradox which this model of belief makes available, and argue that it is more intuitively appealing than the more popular probabilistic solutions.

In chapter 4, I argue that we should extend sensitivism to credences as well as outright belief. In particular, I advance the following two theses:

(CONTEXT) Degrees of belief change from context to context, depending on the space of alternative possibilities.

(UNITY) Outright belief is belief to degree 1.

I claim that (UNITY) solves the usual paradoxes to which threshold views
of outright belief fall prey, and (CONTEXT) undermines the usual reasons
given for rejecting (UNITY).
# Table of Contents

Abstract ......................................................... ii  
Table of Contents ........................................ iv  
List of Figures ............................................... vii  
Acknowledgements ........................................ viii  
Dedication ....................................................... ix  
Introduction ..................................................... 1  

1 Sensitivism ..................................................... 4  
   1.1 Introduction ............................................. 4  
   1.2 Sensitivism ........................................... 5  
   1.3 Precursors ............................................. 8  
   1.4 Lunch with the Boss .................................. 12  
   1.5 Sensitivism and Contextualism ....................... 17  
      1.5.1 Positions on Contextualism May Imply Positions on  
            Sensitivism ....................................... 18  
      1.5.2 Bank Cases ........................................ 23  
   1.6 Conclusion ............................................ 29  

2 Assertion ....................................................... 31  
   2.1 The Knowledge Argument .............................. 33  
      2.1.1 DeRose ........................................... 33  
      2.1.2 Schaffer .......................................... 35  
   2.2 From Knowledge to Belief ............................. 36
List of Figures

1.1 Angelina’s initial state. ................................. 16
1.2 After Brad raises lottery possibilities. .................. 16

2.1 Expresed and asserted content ......................... 51
2.2 Tweety ................................................. 52
2.3 Transitivity .............................................. 62
2.4 Points and their ordering. ............................... 64
2.5 Ordinary and skeptical contexts. ....................... 64

3.1 Is Nixon a pacifist?. .................................. 87
3.2 $S_{pref}$ .................................................. 89

4.1 Credence constraint .................................... 114
Acknowledgements

This dissertation was supported in part by a SSHRC Doctoral Fellowship, and by University Graduate Fellowships from the University of British Columbia. I am grateful for this support.

I presented versions of chapters 2, 3 and 4 at, respectively, the 2010 WCPA at the University of Calgary, the 2011 Notre Dame-Northwestern Graduate Epistemology Conference, and the 2009 Workshop EpiConFor at the Université Nancy 2; I am grateful to all three audiences for valuable comments. I am also grateful to participants in the UBC Philosophy Department’s graduate colloquium, who have provided helpful feedback on nearly every part of this dissertation. Thanks also to Masashi Kasaki, Max Weiss, and my committee (Paul Bartha, Adam Morton, and Patrick Rysiew) for detailed comments on the written material. Above all, thanks to my supervisor, Chris Stephens, for his encouragement, guidance, and painstaking work through every stage of this dissertation’s development.

Needless to say, any remaining errors are my own.
For Becci
Introduction

The purpose of this introduction is to provide some background for non-specialists. In particular, it contains information on the debate over epistemic contextualism, familiarity with which is presumed by chapter 1; all other philosophical debates and problems will be explained in the chapters where they are discussed.

In a sentence, contextualism is the claim that the truth conditions of (utterances of) knowledge-sentences (sentences attributing or denying knowledge to an agent) vary with certain factors of the speaker's context, typically including salience of alternative possibilities and practical stakes. So, for example, a contextualist might say that I could say truly “Jill knows when the next bus will arrive” when Jill and I are standing at the bus stop, engaged in ordinary conversation, but not when Jill and I are in a philosophy seminar where we are seriously discussing the possibility that we are being deceived by an evil demon into thinking there is a next bus, or when it is a matter of life and death that we correctly predict when the next bus will arrive. Non-contextualists might be able to agree with the contextualist’s verdict about this case, for example by claiming that pragmatic differences between the two contexts mean that I only attribute knowledge in one case but not the other. Contextualists, on the other hand, will claim that the difference is due to the context-sensitive semantics of “knows,” though contextualists do not all agree on exactly how the semantics of “knows” is to be spelled out.

Much of the interest in contextualism comes from the often-made claim that contextualism provides an easy solution to skepticism. I will briefly run through the contextualist response to skepticism here as an illustration of

---

1But see Stainton (2010) for an argument that there are plausible views in “the spirit of contextualism” according to which “knows” does not have a context-sensitive semantics.
the uses to which contextualism can be put, although I will almost nothing
to say about skepticism in the body of the dissertation.

Here is an argument for skepticism:

1. I can’t know that I’m not a handless brain in a vat (BIV).
2. If I knew that I had hands, then I’d know that I’m not a BIV.
3. Therefore, I can’t know that I have hands.

This is a skeptical argument because we could easily modify it so that the
conclusion denies any other ordinary knowledge claim. So, as an argument
schema of sorts, this establishes skepticism. Contextualism is supposed to
give us an account both of why this argument is bad and of why it seems
compelling. Briefly, the contextualist response has it that bringing up the
possibility that one is a handless brain in a vat raises the standards for
knowing—now one must have evidence that rules out BIV-possibilities in
order to count as knowing anything incompatible with being a BIV. But
this is not the standard that matters for ordinary knowledge ascriptions,
including most of our ascriptions of knowledge about handlessness. That is,
the contextualist response to the skeptic has it that the argument above does
establish its conclusion, but that conclusion does not conflict with our having
knowledge of ordinary things in the sense we ordinarily think of. We may
not know that we have hands in this inflated, BIV-related sense of knowing,
but that is not a troubling conclusion; we do not ordinarily take ourselves
to have that kind of knowledge.

We have seen that contextualism in epistemology is a view about the
semantics of “know,” in particular the view that truth conditions of “know”-
sentences are sensitive to certain facts about the speaker’s context. The
negation of contextualism is invariantism, which comes in various forms.
Of particular importance to us here is sensitive invariantism (sometimes
called subject-sensitive invariantism, or interest-relative invariantism), as
advocated by Hawthorne (2004), Stanley (2005), and Fantl and McGrath
(2002, 2009). According to sensitive invariantism, there is no special context-
sensitivity in the semantics of the word “know,” but the epistemic standard
one must meet to count as knowing does vary with one’s context: it depends, in particular, on the stakes or practical importance of the proposition in question. Thus, whereas contextualists characteristically think that the truth conditions of knowledge ascriptions are sensitive to the context of the *speaker*, sensitive invariantists characteristically think that they are sensitive to the context of the *subject*.

I will be arguing for a view about belief analogous to sensitive invariantism. That is, I will be arguing that whether one believes that \( p \) depends, in part, on certain facts about one’s context; I claim that belief is subject-sensitive in the same sense that sensitive invariantists claim that knowledge is (though not to exactly the same facts about context), but I make no claim about whether belief ascriptions are speaker-sensitive in the same sense that contextualists claim that knowledge-ascriptions are, though I intend to take up that question in future work. For clarity, because there is sometimes a danger of confusion over what sort of contextualism one is speaking of, I will use the term “k-contextualism” to refer to the standard sort of contextualism described above, and “b-contextualism” to refer to an analogous view about the words we use to ascribe belief.

The reader who wants more extensive background on contextualism and its rivals should refer to Rysiew (2009), a concise and accessible overview of the debate.
Chapter 1

Sensitivism

1.1 Introduction

The aim of this dissertation is to argue for sensitivism about belief. This chapter stands somewhat apart from those to follow: subsequent chapters provide arguments that we should be sensitivists, whereas this chapter argues that we should be interested in sensitivism. The reasons for being interested in sensitivism come in three main categories. First, there are several philosophers (including Kyburg, Levi, Nozick, Stalnaker, and Weatherson) who have argued for views close to sensitivism; anyone interested in the work of these philosophers should be interested in sensitivism as well. In §1.3, I survey the relevant views, and briefly explain where mine connects with and diverges from them. Second, as I argue in the remainder of this chapter, being interested in the debate over contextualism in epistemology is a reason for being interested in whether sensitivism is true. In §1.4, I offer a case in support of sensitivism not unlike those commonly used to motivate contextualism, and in §1.5, I argue that sensitivism provides compelling alternative diagnoses of some of the contextualist’s central cases. Third, the form of sensitivism for which I argue should be of interest to Bayesians and their ilk—roughly, anyone used to dealing with probabilistic representations of partial belief—because I will argue in chapter 4 that sensitivism extends to credences as well as outright belief. For this last remark to make any sense, and to get started on the work of this chapter, we will need a statement of just what sensitivism is; that is the job of §1.2, at the end of which I will return to the credence-related reason for interest in sensitivism.
1.2 Sensitivism

I claim that whether $S$ believes that $p$ depends on $S$’s context.\(^2\) So far, this is nothing new. People do change their minds, after all, so $S$ might believe that $p$ in one context and disbelieve that $p$ in another. To make my claim more specific, then: the facts about context which contextualists\(^3\) say affect the content of knowledge attributions made in that context affect what is believed by agents in that context. To be a bit more precise: contextualists claim that certain properties of the context of the speaker of an attribution or denial of knowledge help determine whether the subject of that attribution or denial knows or fails to know the proposition in question; I claim that the same properties of the context of an agent help determine whether that agent believes or disbelieves a given proposition. In short, the contextualist cares about the attributor’s context, and I care about the agent’s context, but we are both interested in the same properties of the contexts we respectively care about. Other things being equal, making alternative not-$p$ possibilities salient to an agent will make the agent cease to believe that $p$. Other things being equal, raising the practical importance of $p$ for an agent will make the agent cease to believe that $p$. Furthermore, these sorts of shifts across context do not amount to changing one’s mind: this is not best regarded as belief revision. (In §1.4, we will see an intuitive reason to take some such shifts not to amount to changing one’s mind; in chapter 4 we will see how belief revision is handled on my favoured view of belief.)

Because this sort of shiftiness is a general phenomenon, I argue that we should regard beliefs as context-relative. More precisely, the idea is that one and the same underlying mental state may count as belief in one context but not another. Consequently, one can shift from believing that $p$ to disbelieving

\(^2\)Note that content externalism of the sort advanced by, e.g., Kripke (1972), Putnam (1975) and McGinn (1977) can be regarded as claiming that whether $S$ believes that $p$ depends on extrinsic facts about $S$’s context. However, this is not the sort of context-sensitivity I will be concerned with.

\(^3\)Henceforth, unless indicated otherwise, I use “contextualist” to refer to contextualists about “know”, not contextualists about “believe”, of whom there are so far not very many. For clarity, I will occasionally use “b-contextualist” and “k-contextualist” to distinguish the two types of contextualist.
that $p$ without changing one’s mind, in a sense.

In §1.4, I will flesh out in more detail the specific sensitivist model of belief I favour; further detail will be added in chapters 2 and 4, including an account of how outright belief fits with degrees of belief. This relationship is important for understanding my account of outright belief (or binary belief, full belief, belief *simpliciter*, etc.); but for the moment it will suffice to say, as a first approximation, that outright belief supervenes on degrees of belief plus certain aspects of the believer’s context. Thus, holding all of an agent’s degrees of belief fixed but varying the agent’s context may lead to a change in the agent’s outright beliefs.

Note that the question of sensitivism is orthogonal to the question of contextualism about “believe”. That is, taking a position as a sensitivist or non-sensitivist does not commit one to a position in the (so far, largely hypothetical) debate between b-contextualists and non-b-contextualists. This is because, as explained in the Introduction, contextualism has to do with attributor-sensitivity (sentences of the form “$S$ believes that $p$” having their semantics determined in part by the context of the speaker), whereas sensitivism has to do with subject-sensitivity (roughly, whether $S$ believes that $p$ or not being determined in part by $S$’s context). I will not commit myself one way or the other with respect to contextualism about “believe”; although I think the question is as worth investigating as contextualism about “know” is, that is beyond the scope of this dissertation.

I present this dissertation as a work in epistemology. But why is it not just armchair psychology? It sounds, so far, like sensitivism just has to do with how people tend to form beliefs—a question best answered by empirical psychology, surely. I have three responses to this worry. First, some empirical work along these lines has already been done, which partly informs what

---

4Recall from the Introduction: *b-contextualists* are contextualists about words used to ascribe belief, whereas *k-contextualists* are contextualists about “know” and related words.

5For example: Mayseless and Kruglanski (1987) found that people tend to have less subjective confidence in their judgments when stakes are high, and according to Kelley’s (1971) “discounting principle,” people tend to have lower confidence in a given hypothesis when there are more alternative hypotheses available. Both of these observations are discussed in Nagel (2008). Also, see Nagel (2010) for an extended discussion of the “availability heuristic” which fits nicely with the view of belief I advance in the chapters.
I have to say. So this is not purely “armchair” theorizing. Second, psychologists cannot directly observe anyone’s beliefs. They can observe behaviour, judgments, assertions, self-assessments, and the like, but it takes some philosophical work to draw conclusions about what an agent believes from that kind of data. Sensitivism is a thesis about what our concept of belief should be like, and this will influence how we interpret the psychologists’ empirical results. So the question of sensitivism is, at least, a philosophical one. Third, I think the question should be of interest to epistemologists because of the sorts of arguments I will offer for sensitivism in the chapters to follow this one. For the most part, my arguments for adopting sensitivism turn on the philosophical problems that can be easily resolved, given a sensitivist model of belief. In chapter 2, I argue that we need sensitivism to hold on to the familiar principle linking belief and sincere assertion. In chapter 3, I argue that sensitivism allows for a nice resolution of the preface paradox. In chapter 4, I argue that on my preferred sensitivist model of belief, we can neatly reconcile outright belief with degrees of belief, without running into the usual paradoxes. These are all philosophical problems, and the latter two, at least, are clearly epistemological ones.

We are now in a position to say a little more about the credence-related reason for interest in sensitivism. In chapter 4, I will argue that not only outright belief, but also partial belief is sensitive to context in the way I have hinted at above. Here is one way to see that this would require a radical re-thinking of how credences work: one obvious way to change the practical importance of \( p \) for an agent—and, hence, to make a change of context relevant to one’s credences, according to the proposed version of sensitivism—is to offer the agent a bet on \( p \) at very high stakes. It has been my experience that little more needs to be said to Bayesians about why credence-sensitivism is interesting, and so I will defer further discussion of credences to chapter 4.

to follow. However, a full discussion of how my view (and its rivals) would explain the various empirical data is outside the scope of the present work.
1.3 Precursors

In this section, I briefly survey some earlier views which are similar to my sensitivism, and the ways they differ.

Nozick (1993, 96) offers “a (very) tentative suggestion”:

A belief excludes possibilities in a (type of) context. In another type of context, those very possibilities would not be excluded. I believe my new junior colleague is not a child molester. (Asked to list the people in the philosophy building who are not child molesters, I unhesitatingly place his name on the list.) Now the context changes; I need someone to watch my young child for two weeks. A mistake here would be serious—the stakes have escalated. Now I think more carefully. It is not that I did not believe in my colleague’s innocence before. In that context, for those purposes, I did believe it; I did not consider, or assign a probability to, the possibility of his being a child molester. In this context, with higher stakes, I consider what probability that might have.

This sounds very close to the picture of belief I will offer in chapter 2: belief is context-sensitive; in particular, belief that \( p \) is sensitive to the practical importance of \( p \) in the context at hand; and to believe that \( p \) in a context is to exclude not-\( p \) possibilities. But there are some important differences, too. For one thing, Nozick is concerned here with finding a middle path between “the radical Bayesian,” who thinks there are only degrees of belief and no outright beliefs, and a Levi-style view on which belief confusingly implies both certainty and revisability. (I will have more to say about Levi below.) So Nozick’s suggestion here is not quite that outright belief is context-sensitive, nor that degrees of belief are context-sensitive, but rather that one may shift between having outright beliefs about \( p \) and having only degrees of belief about \( p \), depending on contextual factors.\(^6\) In contrast, we will see in

\(^6\)Nozick (1993, 98–9) does seem to think that belief is context-sensitive in something more like the sense I have in mind—for one thing, he speaks of an agent as being disposed
chapter 4 that on my view, outright belief is reducible to degrees of belief. Another important difference between Nozick’s suggestion and my view is that he only definitely identifies high stakes as a contextual factor to which belief is sensitive, although he does speculate that there may be other such factors.

The view advanced in Levi (1980) also shares some important features with my view. Levi’s account is not explicitly about belief—he is primarily concerned with what he calls “knowledge,” though it seems to me that his project is most naturally understood as having to do with rational belief. See in particular Levi (1980, §1.9), which includes this sentence: “From X’s point of view at t, there is no difference between what he fully believes at t and what he knows at t.” Since Levi is concerned with rational belief, whereas I am concerned with mere belief simpliciter, his view and mine are not directly comparable. However, Levi takes rational belief and rational credence to have the following features, which I also claim for ordinary belief and ordinary credence: (rational) credences change from context to context by means other than conditionalization; (rational) full belief coincides with credence 1; and (rational) full belief is the “standard of serious possibility”—i.e., the only possibilities one takes seriously at a given time are those possibilities consistent with one’s beliefs at that time. On the other hand, Levi claims that one’s credences should not be changed without warrant, and in particular should not be changed in response to changes in one’s “goals and values, the problems [one] is investigating, the way [one] has succeeded in identifying potential solutions, and other circumstantial factors” (Levi 1980, 92). Thus, Levi thinks that the rational agent’s credences are not context-sensitive in the way that I claim credences are, in general. This does not mean that Levi’s view of rational credence is inconsistent with my view of ordinary credence—after all, Levi’s view is prescriptive where mine is more descriptive—but it does mean there is some tension.

to believe a proposition \( p \) in a type of context \( C \)—but his discussion leaves it unclear exactly how this is to be spelled out, and what relation this view of belief has to his “(very) tentative suggestion,” quoted above.
Kyburg (1983), Weatherson (2005), Ganson (2008), Sturgeon (2008), and Fantl and McGrath (2009) each argue for a different version of what we might call a “shifting thresholds” view of belief. That is, on such views, to believe \( p \) outright is to have a high enough credence in \( p \)—but what counts as high enough credence for belief varies with context (and perhaps with the choice of \( p \)). In particular, changes in the practical importance of \( p \) will generally change the threshold for belief: if the stakes are higher, then one will need a higher credence in \( p \) to count as believing it outright, \textit{ceteris paribus}. I substantially agree with what these authors have to say about outright belief; their views primarily differ from mine in their treatment of credence as context-insensitive. As we will see in chapter 4, I claim that to believe outright that \( p \) is to give \( p \) credence 1; outright belief is context-sensitive because credences themselves are context-sensitive. In a sense, one can see my view as holding the threshold for outright belief fixed, and treating credence as context-sensitive, whereas Weatherson \textit{et al.} treat credence as fixed, and the threshold as context-sensitive.

There is one more approach to belief which should be mentioned here. On a Jeffrey-style (1992) eliminativist account, all talk of outright belief is, at best, sloppy. Jeffrey’s view is that there is no such thing as outright belief; rather, all belief talk is a (better or worse) approximation of credence talk. I will have little to say directly about eliminativism, but this dissertation as a whole can be seen as an argument against eliminativism. It is surely a cost to have to say that ordinary-sounding talk is nonsense; the best reason for being an eliminativist, as I see it, is that we cannot come up with a coherent account of how outright belief works, in some sense. But it is precisely the goal of this dissertation to give an account of outright belief; so if I succeed in my main goal, I will have given an argument against eliminativism.

Finally, before moving on, I will briefly mention that there are some explicit denials of sensitivism in the literature—so sensitivism should not be entirely uncontroversial. Here is Patrick Maher:

Kyburg holds that rational persons accept all and only those propositions which have a probability greater than some thresh-
old value $p$. The value of $p$ may be different in different contexts. (So belief, in this theory, is context relative.) Within a given context, a person is supposed to treat his beliefs (relative to that context) as certainly true for the purposes of deliberation; however, a belief relative to one context need not be a belief relative to another context, and so need not be treated as certainly true in all contexts. As an account of belief, the major drawback of this theory is the fact that belief seems clearly not to be a context-relative notion. ( Maher 1986, 383n6; emphasis added)

And here is Richard Foley:

Belief ... involves an element of commitment, and commitments are context relative. ... But at least for the purposes of epistemology, we want a notion of belief that is not like this. .... We want a notion such that either you believe a proposition or you don’t. We want to avoid a notion that has you believing it relative to one purpose and yet at the same time not believing it relative to some other purpose. (Foley 1993, 199)

Stalnaker (1984) says some things favourable to my view, but comes down against the context-sensitivity of belief. He thinks that belief is one example of an “acceptance” concept; other examples include assumption, supposition, and presupposition. Acceptance concepts in general are supposed to be context sensitive in a way that beliefs are not:

[W]hat a person accepts can be compartmentalized in a way in which what he believes cannot be. A person may accept something in one context, while rejecting it or suspending judgment in another. There need be no conflict that must be resolved when the difference is noticed, and he need not change his mind when he moves from one context to the other. But something is wrong if I have separate incompatible sets of beliefs for different circumstances. I cannot reasonably believe what I disbelieved yesterday without thinking that yesterday’s belief was mistaken. (Stalnaker (1984, 80–1))
Needless to say, I disagree with all three.

1.4 Lunch with the Boss

Consider the following variations on a dialogue. Angelina hates her job, and she hasn’t been doing very well at it. Tomorrow, she’s got an important meeting with the boss at lunchtime—if she doesn’t go to the meeting, she’s sure she’ll lose her job. But Angelina needs this job—she doesn’t think she can find anything better, so she doesn’t want to get fired. The lottery drawing is tonight.

(LUNCH) Angelina: I’ll be going to lunch with the boss tomorrow.

Brad: Oh, so you believe that your lottery ticket is a loser? After all, if you’re a winner, you won’t be having lunch with the boss tomorrow. Why did you get a ticket if you thought it would lose—or what changed your mind since buying it?

Now there are several ways Angelina could reasonably respond, including each of the following.

(REVISE) Angelina: Oh yeah, I forgot about that. I guess I might not be having lunch with the boss tomorrow after all. Let’s hope not!

OR

(STICK) Angelina: Oh yeah, I forgot about that. Well, still, even if I win the lottery, I’ll be having lunch with the boss tomorrow—I won’t check the draw results until the afternoon.

OR

(ELABORATE) Angelina: Well, come on. If you want to be really pedantic about it, I guess it’s just very likely that I’ll be having lunch with the boss tomorrow. But that’s not what I meant. Of course I’m not sure my ticket will lose—but I realize
that’s most likely. What I meant was that I’ll be having lunch with the boss tomorrow unless something really extraordinary happens. Heck, it’s possible a meteorite could crash into the office tonight, too, and then I certainly wouldn’t be going to lunch with the boss tomorrow.

OR

(REJECT) Angelina: Come on, you know that’s not what I meant. I just meant that I’ll be having lunch with the boss tomorrow unless something really extraordinary happens. But look, let’s be realistic: my ticket isn’t going to win. I’m going to have lunch with the boss tomorrow.

In (LUNCH-REVISE), it seems Angelina has changed her mind over the course of the conversation. At the start, she believed that she would have lunch with the boss tomorrow, but at the end she no longer believes this.

In (LUNCH-STICK), the opposite happens: Angelina sticks to her guns. When the previously-unconsidered lottery possibility comes before her, she still judges that she will have lunch with the boss tomorrow.

In (LUNCH-ELABORATE), as in (LUNCH-REVISE), Angelina offers an opinion which seems to conflict with her initial statement (“it’s just very likely that I’ll be having lunch with the boss tomorrow,” versus “I will be having lunch with the boss tomorrow”). But she also offers an elaboration of her initial claim to specify more precisely what she meant by it, and given this elaboration, there is no conflict. Under this elaboration, Angelina still endorses her initial claim.

In (LUNCH-REJECT), Angelina rejects winning the lottery as a serious possibility. It doesn’t matter where she would have lunch if she won the lottery, because she believes she will not win the lottery. This is the only one of the four versions of (LUNCH) in which Angelina believes outright that her lottery ticket is a loser. However, as in (LUNCH-ELABORATE), Angelina says that her initial claim did not mean what Brad apparently took it to mean: she was, at that point, offering no opinion on whether she would win the lottery, or on where she would eat if she were to win.
All four versions of the (LUNCH) dialogue are natural (at least if construed as conversations over e-mail, to explain why Angelina and Brad are expressing themselves in uninterrupted paragraphs). I wish to offer a model of belief on which we can take what Angelina says at face value, as sincere assertions expressing beliefs. This is not as easy to do on a naïve, context-insensitive model of belief.

To illustrate the difficulty, consider (LUNCH-ELABORATE). If we want to take Angelina’s initial assertion as sincere, then we must take her to believe that she will be having lunch with the boss tomorrow, without qualification. That is, as of her first utterance, she must believe that she will have lunch with the boss tomorrow, regardless of whether she wins the lottery and regardless of whether the office is hit by a meteorite tonight, or else she believes that she will not win the lottery, and that the office will not be hit by a meteorite tonight. But then she must have changed her mind by the time she makes her second utterance: then she claims that it is only very likely that she will have lunch with the boss tomorrow; she does not rule out the possibility of going elsewhere or nowhere for lunch tomorrow, in particular in case she wins the lottery or the office is hit by a meteorite tonight. But Angelina claims not to have changed her mind: what she says about the lottery and meteorite scenarios does not contradict the belief she initially expressed.

On the view I favour, we can take Angelina at her word both before and after Brad’s questions. When she says that she will be having lunch with the boss tomorrow, she believes just that; when she says it’s just very likely that she will be having lunch with the boss tomorrow, she believes just that—and she has not changed her mind.

Here is the idea. At each stage of the conversation, Angelina has a certain range of alternative possibilities in mind. At the beginning of the conversation, in all versions of (LUNCH), no possibility in which Angelina wins the lottery is under consideration. Brad’s questions, however, bring the lottery possibilities to salience, and so Angelina must consider them—if only briefly, as in (LUNCH-REJECT). The four different cases result from different ways Angelina might react to having lottery possibilities made salient.
When Angelina makes her initial assertion, the belief expressed thereby is just that in all the alternative possibilities then under consideration, she will have lunch with the boss tomorrow. In (LUNCH-REJECT), Angelina judges that the lottery possibilities are so unlikely that they are not worth considering. She therefore refuses to give any judgment on where she will have lunch tomorrow in case she wins the lottery. Lottery possibilities have been made salient, but Angelina nevertheless does not take them into consideration; her range of alternative possibilities remains as it was. In (LUNCH-ELABORATE), she concedes that if she wins the lottery, she will not have lunch with the boss tomorrow; she also raises meteorite possibilities to salience and judges that these are also possibilities where she would not have lunch with the boss tomorrow. But because lottery and meteorite possibilities are so unlikely, she still believes, at least, that she is very likely to have lunch with the boss tomorrow. This does not contradict her initial belief, because her initial belief was silent on lottery and meteorite possibilities. In (LUNCH-STICK), Angelina judges that even if she wins the lottery, she will still have lunch with the boss tomorrow. Therefore, even when lottery possibilities are under consideration, she still counts as believing that she will have lunch with the boss tomorrow. We might see this as an expansion of her initial belief: where initially she believed simply that in ordinary (non-lottery) scenarios, she would have lunch with the boss tomorrow, after Brad’s questions, she comes to believe that that she will have lunch with the boss tomorrow in both lottery and ordinary scenarios. Finally, in (LUNCH-REVISE), consideration of lottery possibilities leads Angelina to abandon her initial belief. This may be because she took lottery possibilities to be among her initial range of alternatives, but had overlooked them.

This is all easier to see with pictures.

Think of the rectangle labelled “U” (for “Universe”) in Figure 1.1 as the space of alternatives under consideration by Angelina at the start of the dialogue. Let \( boss \) be the proposition that Angelina will have lunch with the boss tomorrow. The possibilities above the horizontal line in the middle of the box are possibilities where \( boss \) is true; those below the line are possibilities where \( boss \) is false. The possibilities below the horizontal line
are shaded, to indicate that Angelina is ruling them out. That is, according to Angelina, none of the ¬boss possibilities under consideration correspond to the actual world.

Figure 1.1: Angelina’s initial state.

In Figure 1.2, the box \( U \) has been expanded. Now, as well as the ordinary possibilities which were included in \( U \) in Figure 1.1, there are also some lottery possibilities. The proposition that Angelina will win the lottery is represented by \( lott \). The portion of \( U \) where \( lott \) is true is unshaded, to indicate that Angelina is not ruling out the possibility of winning the lottery. Furthermore, there is some overlap between the \( lott \) portion and the ¬boss portion, and this is unshaded: Angelina is not ruling out the possibility of both winning the lottery and missing lunch with the boss—which is why Brad brought up the possibility of winning the lottery.

Figure 1.2: After Brad raises lottery possibilities.

In the state represented by Figure 1.1, Angelina judges that all unrejected possibilities under consideration are ones where she would have lunch with the boss tomorrow; therefore, in that state, she counts as believing that she will have lunch with the boss tomorrow. On the other hand, when Brad asks about what would happen if Angelina wins the lottery, Angelina starts to consider lottery possibilities. If she thereby comes into the state represented
in Figure 1.2, she will no longer count as believing that she will have lunch with the boss tomorrow, because there are now some unrejected possibilities she is considering where she will not have lunch with the boss tomorrow. However, if the ordinary possibilities are much more likely than the lottery possibilities, Angelina will still count as believing that she is very likely to have lunch with the boss tomorrow. If Angelina proportions her belief to her evidence, then Figure 1.2 gives $lott$ far too large a share of the total area of $U$; but, at least, it is much smaller than the area given to $\neg lott$, and this is meant to indicate that Angelina takes $\neg lott$ to be much more likely than $lott$.

1.5 Sensitivism and Contextualism

In this section, I will argue that if you care about epistemic contextualism, you should care about sensitivism too. This is because the question of sensitivism bears on the debate over contextualism in at least the following ways: contextualism nearly implies sensitivism; and sensitivism allows for an invariantist explanation of DeRose’s (1992) bank cases. I take up each of these issues in turn in the following subsections. In §1.5.1, I argue that k-contextualists should be sensitivists; but in §1.5.2, I argue that sensitivism undermines at least some of the arguments for k-contextualism. There may be some tension between these claims, but I am happy with it: if the connections between sensitivism and k-contextualism are tangled, then more philosophers should work on penetrating the tangle and investigating sensitivism—and drumming up interest in sensitivism is one of the main aims of this chapter.

Before we get to the arguments, a disclaimer: I do not mean to present sensitivism as a magic bullet that will, by itself, rule out any of the contenders in the debate over contextualism. I only mean here to make the case that the question of sensitivism is relevant to the question of contextualism, and that anyone interested in the latter question should pay attention to the former as well. Thus, in §1.5.1, I make only the beginnings of an argument that k-contextualists in particular should be sensitivists. The goal of that section
is to highlight a point of contact between contextualism and sensitivism, not to persuade anyone of the truth of sensitivism—that is the business of the three chapters to follow this one. Likewise, in §1.5.2, I argue that sensitivism allows for a nice invariantist treatment of the bank cases, but I do not go on to cover all the other cases contextualists have marshalled in support of their view.\footnote{In particular, third personal cases like Cohen’s (1999) airport cases raise additional complications; but see Bach (2005) and Nagel (2008) for ways of extending the belief-removal approach of §1.5.2 to the airport cases.} Again, my present goal is only to show that sensitivism is relevant to the debate over contextualism, not to argue for or against contextualism.

### 1.5.1 Positions on Contextualism May Imply Positions on Sensitivism

What I aim to show in this subsection is that, at least 
\textit{prima facie}, both k-contextualists and sensitive invariantists should accept sensitivism. I do not quite want to claim that either position \textit{entails} sensitivism—just that they seem to lead to sensitivism. The reason I go no further is because I only aim to give a reason for people interested in k-contextualism and its rivals to be interested in sensitivism: I will show that the falsity of sensitivism would provide at least an argument for classical invariantism, for which reason investigating the truth of sensitivism might make a difference in the debate over k-contextualism. So I am content to establish that it is not obvious that the question of sensitivism is entirely independent of the debate over k-contextualism.

Here’s a hand-waving story about why contextualism makes sense. This should not be confused with an argument for contextualism—rather, it’s something like a just-so story about why we would have developed a language that behaves the way contextualists say ours does.

We talk about knowledge—our own or others\footnote{In particular, third personal cases like Cohen’s (1999) airport cases raise additional complications; but see Bach (2005) and Nagel (2008) for ways of extending the belief-removal approach of §1.5.2 to the airport cases.}—for different reasons. Sometimes we’re concerned with what another person knows because we want to predict what he or she is about to do; sometimes it’s because we want to use that person as a source of our own knowledge; there are even times when it’s because we wonder if anyone else is in a position to refute
the skeptic. Furthermore, each of the purposes I just mentioned for talking about knowledge would amount to different things in different circumstances: for example, if I want to predict whether someone will write in her journal tonight “I saw a tiger today,” I care whether she could tell that it was a tiger rather than a lion; but if I want to predict whether she will run away, all I need is the information that she knew it was a tiger rather than a housecat.\(^8\) Or, if I’m concerned with predicting how someone will answer the question “Do you have hands?”, I might or might not care whether the person can rule out the possibility that she is a brain in a vat, depending on her circumstances. Or it might depend on my own circumstances, if I care how she would answer the question because I want to learn, myself, whether anyone can know such a thing, to escape a skeptical funk of my own. In different cases, we care about different knowledge relations—either because we care whether an agent will be able to discriminate one state of the world from certain others which we might not care about in other cases, or because we care whether the agent can meet a certain evidential standard which we might not care about in other cases. With so many purposes to which we put our talk of “knowledge”, it would be cumbersome to introduce a new verb for each new relation; but at the same time, it’s just false that in every case of talking about knowledge we care about one and the same relation between subject, proposition, and evidence. So we have adapted our language so that the word “know” is flexible enough to denote\(^9\) different relations in different contexts, depending on what we and our interlocutors care about in the context.

I think this sort of story should appeal to contextualists. What other reason is there that our language—indeed, most languages, if contextual-

\(^{8}\)Contrastivists about knowledge such as Adam Morton and Antti Karjalainen (2003; 2008) and Jonathan Schaffer (2004; 2005; 2007; 2008a) are fond of cases like this, but I think they work just as well for the contextualist’s purposes in the current circumstance. That is, because I am offering something like a motivation for contextualism rather than an argument for it, it does not matter if contrastivists might have a better explanation of the cases.

\(^{9}\)This might be where the invariantists get off the bus. A flourishing branch of the debate over contextualism concerns the linguistic evidence that the flexibility in our uses of “know” indicates semantic flexibility. See, e.g., Stanley (2005, esp. ch. 2–3).
ism is anything more than a peculiarity of English and maybe some other Indo-European languages—would have such a word as flexible as “know” is supposed to be, except because we use the word for so many different purposes? But this sort of reasoning carries over to belief (or rather to the words we use to report belief). There are similarly many reasons for talking about our or others’ beliefs. Consider the tiger case above: to predict what the agent will write in her journal, the information that she believes that the animal is a tiger rather than a lion is as useful as the information that she knows it; likewise for the other variant of the case. So, again, we should expect that “believe” can denote different relations in different contexts.

Thus, being a k-contextualist gives a reason to be a b-contextualist. Furthermore, being a b-contextualist gives a reason to be a sensitivist. To be sure, as I have pointed out earlier, b-contextualism and sensitivism are logically independent—one has to do with speaker-sensitivity, the other with subject-sensitivity—but it would be peculiar to think that “belief” has a contextualist semantics but that belief is not sensitive to the believer’s context. For, after all, in the case of first-person belief claims, the speaker and the believer are the same person, so the speaker’s context coincides with the subject’s context. The b-contextualist claims that certain properties of the speaker’s context influence whether the subject counts as believing; the sensitivist claims that the same properties of the subject’s context influence whether the subject counts as believing. So the b-contextualist is at least committed to the following, which is like a metalinguistic version of sensitivism: Whether it would be true to say of oneself that one believes that \( p \) depends on what one thinks the stakes of \( p \) are and on what alternatives to \( p \) are salient to one. Deleting the italicized words in the previous sentence would yield sensitivism. Now, it is of course logically possible to hold that sensitivism is false (that whether one believes that \( p \) does not so depend on context) but that this metalinguistic sensitivism is true (that whether one would express a truth in claiming to believe that \( p \) does so depend on context), but I think it would be a peculiar position to take.

The last move might have been a bit quick.\(^{10}\) There are parallel cases

\(^{10}\)Thanks to Carrie Ichikawa Jenkins for pushing me on this point.
where the analogous view would not be peculiar at all—that is, cases where it makes sense to think both that whether it would be true to say of S that S φs depends on certain features of context, but that whether S φs does not so depend. Whether it would be true to say of oneself that one was born here depends on where one is when one speaks. But whether one was born here does not so depend—it depends on the location of the assessor, not of the subject. Why think that the case for “believe” is different from the case for “born here”? Here is the outline of a story. What the assessor’s context must supply in the case of “born here” is a location, i.e., an argument to plug in to the was-born-at relation. But on the b-contextualist story, it is not plausible that what the speaker’s context must supply is an argument; rather, the context specifies a particular relation, belief-by-standard-X, or belief-given-alternatives-Y. But for general reasons of metaphysical parsimony, we might be more comfortable with a plurality of locations than with a plurality of genuine belief relations. Insofar as the non-sensitivist b-contextualist must postulate just such a plurality of belief relations to account for first-person cases, and insofar as the sensitivist (whether b-contextualist or not) need not so postulate, it would be peculiar to be a non-sensitivist b-contextualist. On the sensitivist story, there is one belief relation, and whether one bears it to a proposition p depends on one’s context. On the non-sensitivist b-contextualist story, there are many such relations, none of which depend on one’s context in the relevant way, and one’s context determines which of these relations is picked out by ascriptions of belief. This gives some prima facie reason for b-contextualists to be sensitivists as well.

This line of argument might raise the following question: if all this is true, why is the situation not parallel in the debate over k-contextualism and sensitive invariantism? Ignoring the fact that the sensitive invariantist is, after all, an invariantist, the interesting thing about sensitive invariantists is that they hold something analogous to sensitivism about knowledge. This is how the sensitive invariantist differs from the classical invariantist. So, to put the question again, why not also say that it would be peculiar to be a k-contextualist without also agreeing with the sensitive invariantists about the subject-sensitivity of knowledge (though, of course, disagreeing
with their invariantism)? In short, the answer is that this is just how things are: it would, indeed, be peculiar to be a k-contextualist who does not believe in the importance of the subject’s context to knowledge. In fact, DeRose (2004, 349) points out that subject-sensitivity is perfectly compatible with his contextualist position; it is no objection to contextualism, he says, that knowledge is sensitive to the subject’s context as well as the speaker’s. But this point is not of very great interest in the debate between k-contextualists and sensitive invariantists, because the dialectical importance of the sensitive invariantist’s postulated subject-sensitivity is that it allows easy explanation of the contextualist’s examples (particularly the Bank Cases, on which see §1.5.2) without committing to a contextualist semantics. The important difference between k-contextualists and sensitive invariantists is along the contextualist–invariantist axis, not along the sensitive–insensitive axis; so it does not much matter that k-contextualists should probably usually be sensitive contextualists rather than insensitive contextualists. But it does matter for our purposes whether b-contextualists should usually be sensitivists, because the axis that matters to this work is the sensitive–insensitive one.

Now consider sensitive invariantists. The sensitive invariantist thinks that, although “know” has an invariant semantics, whether a subject knows some true, believed proposition $p$ depends not only on the subject’s evidence for that proposition, but also on the subject’s practical circumstances. If $p$ is very important to the subject—in particular, if there is some great cost associated with being wrong with respect to $p$—then the subject must meet a higher evidential standard to count as knowing that $p$ than if $p$ were a matter of relative indifference. Just as it would be strange for a k-contextualist to

\[11\]

But note that sensitive invariantists typically do not take intuitions on cases like the Bank Cases to be data, nor do they take their easy explanations of those cases to be particularly strong evidence for their views. Rather, sensitive invariantists usually emphasize the intuitive links between knowledge and action as evidence for sensitive invariantism. However, contextualists do typically take intuitions about the cases to be something like data, or at least a central source of evidence for their views. Thus, the point stands that what is dialectically most important for the sensitive invariantist in arguing against contextualism is that the former can explain away the intuitions the latter appeal to on cases like the Bank Cases, because so doing undermines a central argument for contextualism.
balk at b-contextualism, it would be strange for a sensitive invariantist about knowledge to balk at sensitivism about belief.

According to the sensitive invariantist, knowledge is harder to come by in high stakes scenarios than in low stakes ones. Not only that, but according to most defenders of sensitive invariantism, our intuitions tell us that knowledge is harder to come by in high stakes scenarios. But if our intuitions tell us that we are in a scenario where knowledge that $p$ or that not-$p$ is unusually hard to come by, surely we would be more reluctant to form an outright belief that $p$ or that not-$p$. Sensitive invariantist intuitions about cases where an agent lacks knowledge because of raised stakes (i.e., cases where an agent would count as knowing if the stakes were lower) often have it that the agent lacks knowledge because she should have done more before settling on belief. In the high stakes bank case (see §1.5.2), Hannah lacks knowledge because she needs to rule out more alternatives than she would have to in the low stakes case; since she cannot rule them out, she does not form the belief that the bank will be open on Saturday. It would be peculiar to hold that (a) people generally think knowledge is harder to come by in higher stakes scenarios because more alternatives must then be ruled out, but not that (b) belief is harder to come by in higher stakes scenarios because people wait to rule out more alternatives before settling on belief. If people generally think that knowledge is harder to come by in some particular sort of situation, it is natural to expect that they will be less apt to form settled beliefs in those situations.

1.5.2 Bank Cases

Consider the following two scenarios. They are a version of the so-called “bank cases”, central to the debate over epistemic contextualism.

Low Stakes. Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. It is not important that they do so, as they have no impending bills. But as they drive past the bank, they notice that the lines inside are very long, as
they often are on Friday afternoons. Realizing that it isn’t very important that their paychecks are deposited right away, Hannah says, “I know the bank will be open tomorrow, since I was there just two weeks ago on Saturday morning. So we can deposit our paychecks tomorrow morning.”

*High Stakes.* Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Hannah notes that she was at the bank two weeks before on a Saturday morning, and it was open. But, as Sarah points out, banks do change their hours. Hannah says, “I guess you’re right. I don’t know that the bank will be open tomorrow.”

(Stanley 2005, 3–4; cf. DeRose 1992)

Contextualists claim that, intuitively, Hannah speaks truly at the end of both cases. She is right to say that she knows the bank will be open on Saturday in the low-stakes case, and to say that she does not know the bank will be open in the high-stakes case. Furthermore, contextualism is supposed to be able to explain these intuitions better than any of its rivals.

The contextualist explanation of the bank cases goes roughly like this. There is no difference between the low-stakes and high-stakes cases except for the practical importance to Hannah and Sarah of whether the bank is open on Saturday. In particular, Hannah has exactly the same evidence regarding the bank’s hours in both cases. Therefore, if she lacks knowledge in one case but not the other, this cannot be due to a worsening of her epistemic position, since the practical importance of the proposition under consideration does not affect her epistemic position. Rather, the difference in stakes between the two cases leads to a difference in what relation the word “know” picks out when Hannah claims or denies knowledge in each case: in the low-stakes case, she is talking about a relatively lax knowledge relation, requiring modest evidence; in the high-stakes case, she is talking about a relatively stringent
knowledge relation, requiring a higher standard of evidence. Thus, what she says in both cases is true: her epistemic position (constant across both cases) is good enough to meet the standards required for the lax knowledge relation, so she counts as knowing in the low-stakes case, but it is not good enough to meet the standards required for the stringent knowledge relation, so she counts as not knowing in the high-stakes case.

Bach (2005) gives an invariantist explanation of the bank cases by claiming that in the high-stakes version, Hannah does not believe that the bank will be open, and therefore cannot know. Because the stakes are raised, Bach says, the threshold for belief is also raised: although Hannah has the same level of confidence that the bank will be open in both versions of the case, only in the low-stakes version does that level of confidence suffice for outright belief. Since knowledge requires outright belief, Hannah does not know that the bank is open in the high-stakes version of the case, despite being in no worse an epistemic position than she is in in the low-stakes version.

Note that the picture of belief underlying Bach’s proposal is a sensitivist one: according to Bach, one and the same level of confidence that \( p \) may amount to belief that \( p \) in one context, but not in another. There are, of course, alternative ways of giving a “belief removal” explanation of the bank cases. For example, one might argue that in shifting to the high-stakes version of the case, factors are introduced which lead Hannah to change her mind. That is, whereas Bach has it that Hannah has a constant level of confidence across the low- and high-stakes cases, but the threshold level of confidence required for outright belief changes, this alternative view would have it that the threshold is constant, but Hannah’s level of confidence changes. This seems to be what Hawthorne (2004, 169–170) has in mind in his discussion of a belief removal account; he describes a view on which salience of error “destroys” knowledge by destroying belief.

\[12\] I will have more to say about the threshold view of belief in chapter 4.
Brown (2005, 143–149) targets Bach’s belief removal explanation,\textsuperscript{13} but it seems to me that her arguments do better against the changing-confidence explanation than against Bach’s shifting-thresholds explanation. Here is Brown’s argument, as I understand it.

First, on her reading of Bach, what happens in the high-stakes bank case is that Hannah “has practical reason to raise the level of epistemic justification required for belief” (2005, 144); this is Brown’s reading of what Bach means by Hannah’s threshold for belief being raised. This, I think, is important in making sense of Brown’s response to Bach: if the threshold for belief shifts, it shifts because of something Hannah does, though perhaps unconsciously. According to Brown, Hannah is the one who raises the threshold for belief. With that in mind, Brown points out that in the description of the high-stakes case, there is, of course, no stipulation about whether Hannah believes or disbelieves that the bank will be open. Bach’s explanation of the cases depends on this added detail. To be sure, on the most natural or obvious filling-out of the high-stakes case, it might be that Hannah disbelieves that the bank will be open on Saturday; but the point is that this is a mere psychological detail which could go otherwise. That is, we could have stated the high-stakes case so as to stipulate that Hannah does believe that the bank will be open, by extending the dialogue so that Hannah explicitly affirms her belief that the bank will be open on Saturday, but denies that she knows it, and insists on checking. Brown claims that in this case our intuition that Hannah lacks knowledge would remain.

But it is not clear that Brown’s reading of Bach is the most charitable. That is, Bach is not the only author to propose a shifting-thresholds view of belief (cf., e.g., Weatherson 2005; Ganson 2008; Sturgeon 2008; Fantl and McGrath 2009), but he would be the only one (to my knowledge) to propose that it is the believer who shifts the threshold, rather than impersonal factors brought in by the context. On this understanding of how thresholds shift,

\textsuperscript{13}Brown focusses most explicitly on Bach’s treatment of Cohen’s airport cases rather than DeRose’s bank cases, but this makes no difference for our purposes, because Bach treats both sets of cases similarly, and there is no reason to think Brown’s response to Bach is more compelling for the airport cases than for the bank cases.
Brown’s attack does not fare so well. If raising the threshold for belief is not something Hannah can opt out of doing, or resist doing, in the high-stakes case, then we should expect that even holding her mental state fixed between the low- and high-stakes cases, she may believe in the low-stakes case and disbelieve in the high-stakes case that the bank will be open on Saturday. Of course, the fact that her mental state is held fixed between the two cases is a psychological fact about Hannah which could be otherwise, but if it were different between the two cases—if she had a higher or lower level of confidence in the high-stakes case—then the bank cases would not do the work the contextualist wants them to. After all, the point is supposed to be that everything except the stakes is held fixed between the low- and high-stakes cases; so if there is a difference in what Hannah knows, it must be attributable to the change in stakes. But if there is also a change in Hannah’s level of confidence, that change might be what causes the difference in her knowledge. For example, if she has a higher level of confidence in the high-stakes case despite having exactly the same evidence, then she might not count as knowing because her level of confidence is higher than the evidence warrants. The contextualist (Bach’s target) should want to hold everything other than stakes fixed, including Hannah’s mental state; Bach describes a view of belief on which despite holding Hannah’s mental state fixed across the cases, she nevertheless counts as believing in the low-stakes case and not believing in the high-stakes case. If we accept Bach’s picture of belief, then in Brown’s proposed modification of the high-stakes case to stipulate that Hannah does believe that the bank will be open, Hannah must have a higher level of confidence in that proposition than she does in the low-stakes case.

Fantl and McGrath (2009, 43–45) also argue against belief removal explanations of the bank cases, but they quickly dismiss Bach’s account (along much the same lines as Brown’s response), and focus on Nagel’s (2008) account, which is more empirically informed than Bach’s. Nagel points to results from empirical psychology showing that, in general, people reason more accurately, but more cautiously in high-stakes circumstances than in low-stakes circumstances. As a result, it generally takes more evidence to produce a judgment (belief) that $p$ in a high-stakes context than in a low-
stakes context. Thus, “it’s only psychologically plausible that High-Stakes Hannah could have high confidence of the sort stipulated on her slim evidence if she is compromised in her accuracy, for example by thinking hastily or in a way biased by wishful thinking” (Nagel 2008, 291–2). Therefore, even with the stipulation that Hannah believes the bank will be open on Saturday in the high-stakes case, we should judge that she does not know, on the grounds that she must have come by that belief illicitly.

Fantl and McGrath reply by redescribing the high-stakes bank case so that Hannah is explicitly described as using the same methods as in the low-stakes case to come by a belief that the bank will be open on Saturday:

With equanimity—no more hastily and with no more bias than [she] does in [the low-stakes case], and using exactly the same methods—[Hannah] forms the belief that the bank is open tomorrow. [She] acts on that belief by waiting until the following day, to [her] wife’s amazement and frustration. Of course, [she’s] irrational in so waiting, for the risk is too great and the cost of going today is so small. But [she] nonetheless, without a second thought, does it. We still have the intuition that [she] speaks falsely when [she] says “I know”.

(Fantl and McGrath 2009, 45)

Thus, Fantl and McGrath try to get around the psychological generalizations Nagel brings up by explicitly describing Hannah as violating them in the high-stakes case: regardless of what people generally do in high-stakes situations, in this high-stakes situation, this person reasons just as she would in the corresponding low-stakes case. Since the intuition that Hannah does not know the bank will be open remains, it must not be attributable to lack of belief or lack of well-formed belief in the high-stakes case.

But it is not clear that Fantl and McGrath’s stipulation has preserved what matters about the high-stakes bank case. In particular, the Hannah they describe sounds like someone who simply does not appreciate that the stakes are as high as they are, or does not care, which amounts to the same thing. She is just as nonchalant about the bank’s hours as she would be in
the low-stakes case, and this is hard to understand unless she takes herself to be in the low-stakes case, for all intents and purposes. Of course, Fantl and McGrath could also add a further stipulation that Hannah does understand what is at stake, but once all these stipulations have been made, I find the case completely unintelligible. It is not clear what can be learned from intuitions about such a bizarre case.

Consider the following scenario. Dudley sees that there is a runaway trolley about to run over and kill five people who have been tied to the tracks before the trolley. Dudley also sees that he can stop the trolley, and save the five people, by pushing a very large sleeping football player onto the tracks, and that this is the only way anyone can save the five people—but the football player would certainly die in the process. Dudley thinks that it would be morally wrong to treat the football player as a mere means, however, and thinks that pushing the football player onto the tracks to save the five people (by dying) would be treating the football player as a mere means. Dudley therefore thinks that the morally right thing to do is to leave the football player alone, and let the trolley run over the five people tied to the tracks. Dudley wants to do the morally right thing (and not the morally wrong thing). Therefore, Dudley chooses to push the football player onto the tracks.

Dudley’s psychology is bizarre and (at least to me) unintelligible. Even if facts about Dudley’s psychology might make a difference to some philosophical theory, intuitions about his case should not be trusted, or treated as evidence or data. I think the same is true of the case where Hannah understands what is at stake, but acts contrary to that understanding.

What’s the upshot of all this? A sensitivist view of belief gives a nice analysis of the bank cases, without appealing to anything special about the semantics of “know”.

1.6 Conclusion

By now, it should be clear that sensitivism is at least prima facie plausible, thanks to the Lunch cases presented in §1.4. Also, in §1.5 I argued that
sensitivism should be of interest to participants in the debate over epistemic
contextualism, since it seems contextualists should be sensitivists, and since
a sensitivist view of belief allows an invariantist explanation of (at least) the
widely-discussed bank cases. But I have not yet offered an argument that
sensitivism must be correct; that is the business of chapter 2, which will
examine the connection between belief and assertion.
Chapter 2

Assertion

Introduction

In chapter 1, we saw some reasons to think sensitivism is plausible, and that philosophers interested in epistemic contextualism should be interested in sensitivism. In this chapter, I will pursue similar aims—arguing for sensitivism and for contextualists’ interest therein. In §2.4, I present an argument for sensitivism based on facts about assertion, in particular the thesis that an assertion is sincere if and only if the assertor believes what she asserts. This, I think, is more compelling than the arguments I offered in chapter 1 for sensitivism: the conclusion here is not merely that sensitivism is plausible, but rather that we ought to be sensitivists. In §2.5, I offer a preliminary model of belief based on the considerations raised in giving the assertion-based argument of §2.4.

But before I get to the argument I think is right, I will discuss an argument I think is wrong; this is where we get this chapter’s connection to contextualism. My argument bears some similarity to arguments from DeRose (2002, 2009, ch. 3) and Schaffer (2008b). DeRose and Schaffer both argue from facts about assertion to a version of contextualism. More specifically, they both argue from the knowledge account of assertion (see Unger 1975 and Williamson 2000), and the fact that whether one can warrantedly assert a given proposition \( p \) varies across linguistic contexts, to the conclusion that whether one counts as knowing that \( p \) depends on facts about one’s linguistic context, as well as the usual facts about one’s evidence, the truth of \( p \), etc.

14I say the model is preliminary because it deals only with outright belief. I will not offer a model dealing with degrees of belief until chapter 4; that model will be a generalization of the preliminary one offered here.
The odd thing about these arguments—especially DeRose’s, since he is a
card-carrying contextualist—is that they establish that the truth-conditions
of utterances of the form “S knows that \( p \)” are sensitive to the subject’s (S’s)
context, not the speaker’s context, whereas epistemic contextualism is the
thesis that such utterances are sensitive to the speaker’s context. (This is
why I say that DeRose and Schaffer argue for “a version” of contextualism; cf.
Blackson 2004.) The similarity my argument bears to the DeRose/Schaffer
arguments is that I argue from the fact that whether one can sincerely (not:
warrantedly) assert a given proposition \( p \) varies across linguistic contexts,
and a certain link between assertion and belief, to contextualism about be-
lief. But the important difference is that the link I claim between assertion
and belief is much less controversial than the knowledge account of assertion,
the link DeRose and Schaffer appeal to between assertion and knowledge.
So although there are now several criticisms of the DeRose/Schaffer
arguments in the literature (e.g., Blackson 2004; Hawthorne 2004; Levin
2008; McHugh forthcoming; Shieber 2009), since these criticisms turn on the
knowledge account of assertion, they will not threaten my argument.

I have said that I think the DeRose/Schaffer arguments are wrong. This
is what I will argue in §§2.1–2.3. §2.1 sets out DeRose and Schaffer’s ar-
guments, then §2.2 argues that if you accept the DeRose/Schaffer argument
for contextualism, then you should also accept a cognate argument for sensitivism which results basically by substituting “belief” for “knowledge” throughout their argument. I do not satisfy the antecedent of this conditional, in part because I am not convinced that the knowledge account of assertion is correct, and in part for the reasons I set out in §2.3. There, I argue that the DeRose/Schaffer arguments are best understood as inferences to the best explanation, and that sensitivism provides a better explanation of their explananda than contextualism does. That is, I argue that sensitivism does explain their explananda, and since, according to §2.2, commitment to the DeRose/Schaffer argument requires commitment to sensitivism, my sensitivism explanation is at least as parsimonious as their contextualist ex-

\[15\] Hawthorne does not specifically discuss the DeRose/Schaffer argument, but he does argue that contextualists cannot accept the knowledge account of assertion (2004, 88ff.).
planation. Thus, we have another reason for philosophers who care about contextualism to care about sensitivism too: sensitivism undercuts a prominent argument for contextualism.

2.1 The Knowledge Argument

2.1.1 DeRose

Chapter 3 of DeRose (2009) is an updated and expanded version of DeRose (2002), where the argument from assertion first appeared; I will take the former as my main source in discussing DeRose’s version of the argument, since it represents his considered opinion.

It is clearer in the later work what form DeRose intends the argument to take: the main claim of the chapter is that the contextualist can use a contextualized version of the knowledge norm on assertion to defeat an otherwise formidable objection. DeRose’s target is what he calls “the Generality Objection”: whereas contextualists claim that variation in assertability of knowledge-sentences (attributions and denials) shows that there is variation in truth conditions of knowledge-sentences, the Generality Objector claims that this variation in assertability holds for all sentences. In contexts where the standards for appropriately asserting knowledge-sentences go up, so do the standards for asserting anything else. Therefore, there is nothing special about knowledge-sentences: the variation in their assertability across contexts is no reason to think their truth conditions also so vary.

There is an important dialectical point to be made here. Because DeRose

16 Just what is “assertability,” here? Well, p is assertable for S just in case it is permissible for S to assert that p, or S may assert that p, or S can warrantedly assert that p, or something along those lines. But what sort of “may” is this, what sort of warrant? It is not the “may” of moral or rational permissibility, certainly. Proponents of the knowledge account of assertion often speak of the assertion as something analogous to a game; the knowledge norm is to be understood as standing in a relationship to the practice of assertion analogous to the relationship between the rules of chess and the game itself (Williamson 2000, 238ff.). I am not sure what to make of this, but I will put to one side worries about exactly what kind of assertability we are talking about, because I do not endorse the knowledge account of assertion, and because my main argument, in §2.4, does not rely on assertability. Thanks to Adam Morton for pressing me on this point.
takes his opponent to be arguing from the premise that the epistemic standards one must meet in order to appropriately assert that \( p \) vary across contexts, he does not argue for that premise. (Of course, he does not dispute the premise either; in more than one place, he says it is “obviously” true.) As we will see in §2.3, it makes a difference whether the variation in assertability is really due to a shift in epistemic standards, so it is a little disappointing not to have an argument here.

DeRose’s response to the Generality Objection is that it gets some of the conversational phenomena wrong, while a contextualized version of the knowledge norm on assertion gets those phenomena right. It is not important for our purposes what the details are of DeRose’s response, however, so I omit discussion of them.

But there is another argument for contextualism from the knowledge norm in DeRose’s essay, closer to the sort I will primarily be addressing in this paper. That is, he argues that, given the variation across contexts of the epistemic standards one must meet to assert properly that \( p \), an invariantist understanding of the knowledge norm is untenable: the knowledge account of assertion is “simply incredible” without “contextually variable epistemic standards for the truth of ‘I know that \( p \)’” (2009, 101). We can sum up the argument here as follows:

1. A speaker, \( S \), is well-enough positioned with respect to \( p \) to be able to assert properly that \( p \) if and only if \( S \) knows that \( p \) according to the standards for knowledge that are in place as \( S \) makes her assertion.\(^{17}\)

2. The epistemic standards one must meet to assert properly that \( p \) vary across linguistic contexts.

3. Therefore, whether one knows that \( p \) varies across linguistic contexts.

I think this is how DeRose intends his argument to be read. But, as I will argue in §2.2.3, the argument is better understood differently, as an inference to the best explanation. I will leave further discussion until then.

\(^{17}\)This is DeRose’s principle KAA-R (2009, 99). Contextualists and invariantists will take the principle differently: the latter will think the last clause (about standards for knowledge) does no work, as the same standards are in place in all contexts.
2.1.2 Schaffer

Schaffer’s argument takes a different form than DeRose’s. As I read it, the argument is an inference to the best explanation. There is also a difference in the conclusion Schaffer draws: unlike DeRose, he claims knowledge is a “question-relative” relation. Schaffer begins by arguing for the following thesis about assertion:

(AQ): Whether one may assert that \( p \) in context \( c \) depends on the question under discussion in \( c \). \hspace{1cm} \text{(Schaffer 2008b, 8)}

(“The question under discussion” is taken to be part of the linguistic context at any given point in a conversation.) Schaffer then goes through the standard Unger-Williamson motivations for the knowledge account of assertion, but argues that the standard form of the knowledge norm on assertion does not make adequate sense of AQ. That is, the truth of AQ is not explained by KA:

(KA): \( S \) ought: assert that \( p \) only if \([S]\) knows that \( p \).

On the other hand, the following alternative formulation of the knowledge norm on assertion does explain the truth of AQ:

(KQ): \( S \) ought: assert that \( p \) in context \( c \) only if \([S]\) knows the answer (\( p \)) to the question under discussion in \( c \).

\hspace{1cm} \text{(Schaffer 2008b, 10)}

Therefore, KQ is to be preferred over KA, because it explains more. Schaffer’s argument is not that KA is inconsistent with AQ, or that if KA is true AQ cannot be explained—some other norm on assertion might explain AQ. The argument is, rather, that we ought to prefer a norm which both satisfies the standard Unger-Williamson considerations and also explains AQ over a norm which only does the former.

The upshot of all this is that, since the knowledge relation in KQ is a ternary relation between subject, proposition, and question under discussion, and since the question under discussion is a context-sensitive factor, we have established (some form of) contextualism.

\footnote{Cf. Unger (1975, 250–71) and Williamson (2000, 238–69).}
2.2 From Knowledge to Belief

I have two main claims to make about the arguments just presented. First—the thesis of this section—anyone who accepts these arguments should also accept the result of substituting “belief” for “knowledge” in their conclusions. Second, the arguments above are vulnerable to certain objections which do not touch the analogous arguments I will give in §2.4 for the conclusion that belief is sensitive to context in just the way that DeRose and Schaffer claim knowledge is.

2.2.1 Deriving a Belief Norm from the Knowledge Norm

The crucial insight of this section is that we can derive from the knowledge norm on assertion a corresponding belief norm. Because there are a few different ways the knowledge norm might be formulated, we will have to derive the belief norm in a few different ways. (I defer to §2.2.3 discussion of biconditional forms of the norm, such as the one in DeRose’s argument.)

Here’s one possible formulation, probably the easiest from which to derive a belief norm:

\[(\text{NEC-KA}) \text{ Necessarily, if one may assert that } p, \text{ then one knows that } p.\]

Provided that knowledge necessarily implies belief—which is almost uncontroversial\(^{20}\)—it follows that:

\[(\text{NEC-BA}) \text{ Necessarily, if one may assert that } p, \text{ then one believes that } p.\]

The inference here is valid in any normal modal logic. No strong assumptions about necessity are required.

\(^{19}\)The “may” here is meant to be the “may” of assertability—that is, S may assert that \(p\) in this sense just in case \(p\) is assertable for S. Cf. note 16 on page 33. Likewise, the “ought” in the principles to follow is the dual of this “may”—one may \(\phi\) if and only if it is not the case that one ought not to \(\phi\).

\(^{20}\)But see Radford (1969).
But, for one reason or another, (NEC-KA) is not the most common formulation of the knowledge norm among proponents of the knowledge argument.  More common is this, found in Schaffer (2008b, 9) and Williamson (2000, 243):

(KA): One ought: Assert $p$ only if one knows that $p$.

With this formulation, we do not need as strong a connection between knowledge and belief as we did to move from (NEC-K) to (NEC-B). Even if it is possible that one know that $p$ without believing that $p$, so long as one ought to believe all that one knows, we can derive a belief norm on assertion from (KA). That is, given (KA) and (BK), (BA) follows.

(BK): One ought: Believe $p$ if one knows that $p$.

(BA): One ought: Assert $p$ only if one believes that $p$.

The inference here presumes that “ought” behaves as a normal modal operator analogous to necessity in each of the three principles, and that the “if”s and “only if”s are material conditionals. Again, the inference is valid in any normal modal (deontic) logic.

(BK), I take it, should be acceptable even to those who think that one can know things one does not believe. Even if, for example, excessive modesty or self-doubt prevents one from quite believing something one actually knows, it should be clear that this is a poor state to be in. If one knows that $p$, there is nothing wrong with believing that $p$; and without believing that $p$ one cannot act on one’s knowledge.

One might worry, though, that something fishy has happened along the way from (KA) to (BA). Are we sure those are all the same “ought”s? Are we sure they behave like normal modal operators? These might be hard questions, so I’ll try to avoid them. For those who think a belief norm on assertion is not directly derivable from the knowledge norm, I’ll argue that a belief norm is motivated by the same factors that motivate the knowledge

\footnote{DeRose’s KAA-R can be read as a strengthening of (NEC-KA). Provided that $S$’s being well-enough positioned with respect to $p$ to assert that $p$ follows from the fact that $S$ may assert that $p$, (NEC-KA) follows from the left-to-right direction of KAA-R.}
norm. That is, for each of the standard arguments for the knowledge norm on assertion, there is an analogous argument for a belief norm on assertion. However, I do not claim that a belief norm will explain all the phenomena listed by proponents of the knowledge norm: the point here is not that a belief norm would be an adequate replacement for the knowledge norm, but rather that if you want a knowledge norm, you should want a belief norm too.

Schaffer (2008b, 8) helpfully distinguishes four points used by Unger (1975) and Williamson (2000) to argue for the knowledge norm: Critique, Defect, Challenge, and Authority. (DeRose, for the most part, does not explicitly argue for the knowledge norm, except to discuss apparent counterexamples from Weiner 2005; see 2009, §5.) I will consider each in turn.

**Critique:** One who asserts that \( p \) but does not know that \( p \) is liable to criticism.

Likewise, one who asserts that \( p \) but does not believe that \( p \) is liable to criticism. Asserting what one does not believe is surely even worse than asserting what one believes but does not know—at least the latter assertor is sincere.

**Defect:** Assertions of the form: “\( p \), and I do not know that \( p \)” are defective.

Likewise, assertions of the form: “\( p \), and I do not believe that \( p \)” are defective. These are just two different versions of Moore’s paradox. Remember that for present purposes, I do not need a belief norm on assertion to be able to explain the knowledge version of Moore’s paradox. I only need that there also be a belief version of the paradox.

**Challenge:** An assertion that \( p \) may be challenged by “How do you know” questions.

Similarly, one might challenge an assertion that \( p \) by asking, “Why do you think that?” (Or, more strongly, “Why would you think that?”) This sort of challenge will, in fact, be more natural than “How do you know”-style
challenges in cases where $p$ is something the challenger takes to be obviously false: “Why would you think the Earth is six thousand years old?”

**Authority:** Asserting, like commanding, requires authority.

The authority required for assertion is epistemic authority. It’s hard to come by better epistemic authority than knowledge, so this is a point in favour of the knowledge norm on assertion. Merely believing what one asserts does not confer any comparable authority on an assertor. However, lacking belief that $p$ generally undermines one’s authority to assert that $p$. Indeed, any authority an assertor might have which is compatible with lack of belief must also be compatible with lack of knowledge. Therefore, to the extent that Authority is a reason to favour the knowledge norm, it is a reason to favour the belief norm as well. After all, the cases our proposed belief norm would rule out are those where one asserts something one does not believe.

Finally, note that all the arguments above work just as well for Schaffer’s question-relative version of the knowledge norm as they do for the simpler version. That is, if you think that (BA) follows from (KA) and (BK), then you should also think that (BAQ) follows from (KAQ) and (BKQ):

(KAQ): One ought: assert that $p$ in context $c$ only if one knows the answer ($p$) to the question under discussion in $c$.

(BKQ): One ought: where $p$ is the answer to the question under discussion in context $c$, believe that $p$, if one knows that $p$.

(BAQ): One ought: assert that $p$ in context $c$ only if one believes that $p$, where $p$ is the answer to the question under discussion in context $c$.

---

22 There are, of course, exceptions—cases where a speaker lacks belief that $p$ but still has the authority to assert that $p$. Cf., e.g., the creationist schoolteacher case discussed in Lackey (1999) and Graham (2000): the creationist schoolteacher believes evolution is false, but conscientiously teaches her students exactly what the textbooks say about it. When she unearths a fossilized dinosaur bone in her backyard, she has the authority to assert to her students that it is millions of years old, even though she does not believe it. However, my point here is merely that lack of belief generally undermines one’s authority to assert, not that it always does so; nor can any more be said for knowledge. Notably, Lackey and Graham discuss the creationist schoolteacher as a case where the speaker lacks knowledge as well as belief.
Likewise for deriving (NEC-BA) from (NEC-KA). And if one takes Critique, Defect, Challenge, and Authority to motivate (KAQ), one ought to take them to motivate (BAQ) as well. The arguments I gave above that these four points motivate (BA) if they motivate (KA) also establish the point for the question-relative versions of these norms. Of course, whether one should prefer (BAQ) or (BA) is another question, which we will address shortly.

2.2.2 Schaffer’s Argument Applied to Belief

So let us take it for granted that proponents of the knowledge norm on assertion should also accept a corresponding belief norm on assertion. What follows about the context-sensitivity of belief?

First, consider the reasons Schaffer gives for being dissatisfied with (KA) as opposed to (KAQ). (KA), Schaffer says, has two interrelated problems: it is “blind to the act” (2008b, 9); and it fails to explain (AQ). That is, the norm (KA) only concerns the content of what is asserted, ignoring what the effects of the assertion are. “The norms for assertion should be licenses to do what assertion does,” according to Schaffer (2008b, 9). Second, since (KA) is insensitive to the question under discussion in a context, it cannot explain the fact that whether one may assert that $p$ depends on whether $p$ answers the question under discussion.

(KAQ) suffers from neither of these flaws. First, according to Schaffer, what one does in asserting is aim to answer the question under discussion. Therefore, since (KAQ) only gives one license to assert when one knows the answer to the question under discussion, (KAQ) speaks to what assertion does. It gives license to assert $p$ when one knows that asserting $p$ will do what assertion aims to do. Second, clearly, (KAQ) can explain (AQ): whether one may assert that $p$ depends on the question under discussion, because whether one knows the answer to the question under discussion depends on what the question is. For these reasons, Schaffer argues that (KAQ) is preferable to (KA). Since the knowledge relation in (KAQ) is question-relative, we ought to consider knowledge to be question-sensitive, and hence sensitive to the subject’s linguistic context.
I claim that analogous points hold for (BA) and (BAQ). (BA) cannot explain (AQ), and is blind to the act in the same sense that (KA) is. But the licenses that follow from (BAQ) depend on the question under discussion, which explains (AQ); and since those licenses concern answering the question under discussion, they are licenses to do what assertion does. If one believes that $p$, where $p$ answers the question under discussion, then one believes that asserting $p$ accomplishes the goal of assertion. I conclude that if one accepts Schaffer’s argument for question-sensitivity of knowledge, then one ought also to accept the question-sensitivity of belief.

### 2.2.3 DeRose’s Argument Reconsidered

Turning to DeRose, recall the argument we attributed to him in §2.1.1:

1. A speaker, $S$, is well-enough positioned with respect to $p$ to be able assert properly that $p$ if and only if $S$ knows that $p$ according to the standards for knowledge that are in place as $S$ makes her assertion.

2. The epistemic standards one must meet to assert properly that $p$ vary across linguistic contexts.

3. Therefore, whether one knows that $p$ varies across linguistic contexts.

As I will argue presently, both directions of the biconditional in (1) are needed to make the argument go through. However, I have given no reason to accept an analogue of the right-to-left direction for belief. Therefore, if one understands DeRose’s argument thus, there is no analogue for belief to which one would be committed. However, I think DeRose’s argument is more successful as an inference to the best explanation, closer to the form of Schaffer’s argument. To be clear: I think that DeRose intends his argument to be read as I have it here; what I will now argue is that the argument is less vulnerable to objections if understood more along Schaffer’s lines. If this is correct, then the argument I gave above dealing with Schaffer’s argument would hold for DeRose’s as well.

The trouble here, as pointed out by Shieber (2009, 174–5), is that DeRose does not offer any argument for the right-to-left direction of (1), nor does
Williamson in the piece DeRose cites—but both directions are needed for the argument as it stands. For suppose we have two cases, A and B, which differ only in the relevant linguistic contexts (and these contexts differ only in the ways of interest to contextualists). Suppose further that for the protagonist of both cases, S, some proposition p is properly assertable in A, but is not properly assertable in B because S’s epistemic position is not strong enough with respect to p. Then from the left-to-right direction of (1), we have that S knows that p in A; and from the right-to-left direction of (1), we have that S does not know that p in B. This is how I understand the argument to work; if I am right, then it makes essential use of both directions of (1).

Furthermore, (2) may not be as obvious as DeRose thinks. Of course, I concede that it is obvious that whether one can properly assert that p varies across linguistic contexts, but I do not think it is obvious that this variation in assertability is traceable to shifting epistemic standards. I do not even mean to contend here that (2) is false, just that it is non-obvious; it requires an argument. And I think any plausible argument is likely to be an inference to the best explanation. This leads me to my point about how we should understand DeRose’s argument: we should regard it as an inference to the best explanation, where the data to be explained are both the phenomena motivating the knowledge norm and the cases that motivate (2).

The advantage of understanding DeRose’s argument thus is that it needs only one direction of the biconditional in (1)—as we saw with Schaffer’s argument. Furthermore, the cases for the knowledge norm and for (2) are bolstered: the phenomena each are supposed to explain are now given a unified explanation. A context-sensitive knowledge norm explains both why epistemic standards for assertability vary across contexts (because one must know to assert properly, and knowledge depends on context) and explains \textit{Critique, Defect, Challenge,} and \textit{Authority.}

Now, if I am right that DeRose’s argument is stronger if understood this way, then the right-to-left direction of (1) is dispensable. And by the arguments I gave at the start of this section, something like (BA) follows from the left-to-right direction of (1). This, as I have already argued while considering Schaffer’s version of the argument from assertion, suffices to es-
tablish the context-sensitivity of belief, given (2). In fact, putting aside the differences between question-sensitivity and context-sensitivity, the understanding of DeRose’s argument I have been urging here makes it exactly the same as Schaffer’s.

2.3 Problems with the Knowledge Argument

But there are reasons not to accept DeRose and Schaffer’s arguments. I give some of those reasons here, and in the next section offer an argument for context-sensitivity of belief that avoids these problems. To be perfectly clear, though, my point here is not to argue that DeRose and Schaffer’s arguments should be rejected. Rather, I claim that there are objections they must answer which the argument I will give in the next section need not answer.

First and most obviously, the knowledge norm on assertion is itself controversial. Weiner (2005) and Lackey (2007) both argue that there are cases where one may assert what one does not know; see also Brown (2006, 2008), Douven (2006), and Hindriks (2007). Schaffer (2008b, n5) invites the reader to take his thesis as a conditional: “if knowledge is a norm of assertion, then knowledge must be a question-relative relation.” Like Schaffer, I do not wish to evaluate the objections to the knowledge norm here; instead I note that there are those who object, and will offer an argument for context-sensitivity of belief which I think will satisfy even them.

Second, as I claimed at the end of the previous section, it is not perfectly obvious that the variation in assertability across contexts is due to variation in epistemic standards. To put it another way, even granted the knowledge norm on assertion—even granted that the knowledge norm is the unique norm specific to assertion, as Williamson would have it—it does not follow that every case where assertion is not warranted is a case where the knowledge norm is violated. To take a trivial example, suppose I know that Santa Claus does not exist. I may be warranted in asserting that Santa Claus does not exist in a context where the question under discussion is: “What are some examples of non-referring proper names for my lecture tomorrow?”
I may not be warranted in asserting that Santa Claus does not exist in a context where the question under discussion is: “What is Santa going to bring this year?”. (We may suppose that in both cases, I am talking to my brother, while my infant niece is in the room.) The difference in assertability does not reveal a difference in knowledge.

The moral here is that although knowledge may be required for warranted assertion, there are lots of reasons an assertion can lack warrant. Some of these reasons may be sensitive to linguistic context but nevertheless not be derived from norms specific to assertion. In the Santa Claus cases above, the difference in warrant might come from some norm on whether one should remove pleasant delusions from other people’s children—and such removal can be effected without asserting anything.

Consider DeRose’s bank cases, which he uses to argue for contextualism. In both cases, it is Friday, and the main characters have a cheque to deposit. It would be inconvenient to deposit it right now, so one character asks the other whether the bank will be open on Saturday. The other character has been to the bank on a Saturday a few weeks ago, but has not checked to see if the bank’s hours have changed in the meantime. In the “low standards” case, it is not important to the characters whether they deposit the cheque promptly; in the “high standards” case, they need to have the money deposited before the start of business on Monday. DeRose claims—and I agree—that in the “low standards” case, it would be appropriate to answer the question by asserting that the bank is open on Saturdays, but it would be inappropriate to do so in the “high standards” case. (After all, banks do sometimes change their hours.) DeRose claims that the difference in appropriateness is because, although in both cases the characters’ epistemic position is the same with respect to the asserted proposition, in the “high standards” case the standards one’s epistemic position must meet to amount to knowledge are higher than in the “low standards” case.

To be sure, the knowledge norm on assertion together with context-sensitivity of knowledge would explain the difference in assertability between the two cases. The trouble is that it’s not the only explanation. It might be that some other, more universal norm overrides the assertion-specific knowl-
edge norm in this case. For example, it might be that the assertor in the “high standards” case does satisfy the knowledge norm (that is, she does know that the bank is open on Saturday), but violates the more general norm, “Do not lead others to act in ways you know are imprudent for them.” And, indeed, it is a point of contention between contextualists (like DeRose) and classical invariantists who accept the knowledge norm (like Williamson) how cases like these should be understood.\footnote{A similar point to the preceding is made in Shieber (2009, 174–7); but Shieber does not make the move to understand DeRose’s argument as an inference to the best explanation.}

Finally, I have argued that both Schaffer’s argument and DeRose’s should be understood as arguments to the best explanation. But we may already have a better explanation on hand: what I have argued in the previous section is that anyone who accepts DeRose and Schaffer’s arguments for contextualism about knowledge should also accept sensitivism about belief. But sensitivism about belief explains all the same cases that contextualism about knowledge (of the sort under discussion in this chapter) is supposed to explain. The important point here, as pointed out by Blackson (2004), is that what the DeRose/Schaffer arguments establish is only a first-personal contextualism: we have that the truth of first-person knowledge attributions depends on linguistic context. Therefore, what we would need to show that knowledge-contextualism explains what belief-sensitivism cannot is a case where, clearly, a speaker ought not to assert something she believes, because she does not know it. None of the cases produced by Schaffer or DeRose fit the bill. See §1.5.2 for an argument that no such case will be forthcoming—that sensitivism can explain any first-personal case explained by knowledge-contextualism. In particular, note that in DeRose’s “high standards” bank case, it is intuitive that because of the high stakes involved, the protagonist is too cautious to count as believing that the bank will be open on Saturday (cf. §1.5.2).

So not only do the arguments for context-sensitivity of knowledge rely on a defeasible implication, we may have already found a defeater. I think we can do better in an argument for context-sensitivity of belief, and I attempt to do so in the next section.
2.4 The Belief Argument

The link between assertion and belief goes through the notion of sincerity: one asserts sincerely if and only if one believes what one asserts. Again, slightly differently: one’s assertion that $p$ is sincere if and only if one believes $p$. I will argue that we ought to take the content of an assertion that $p$—and hence the content of the belief one must have for such an assertion to be sincere—to depend on context.

To begin, I will make a distinction: between the content expressed by an assertive utterance and the content asserted by such an utterance. Roughly, the guiding idea behind the distinction here is for asserted content to be the sincerity-relevant stuff—whatever needs to be mirrored in the assertor’s beliefs in order for her assertion to count as sincere—whereas expressed content is something closer to ordinary speaker-meaning or utterance-meaning. I will have little to say about expressed content except in its relationship to and contrast with asserted content. In general, the expressed content of an assertive utterance, plus some facts about the speaker’s intentions and beliefs about the context, will determine the utterance’s asserted content. The following two subsections are devoted to exploring the distinction between expressed and asserted content. Because of the guiding idea behind asserted content, I will begin with some observations about sincerity.

Note, incidentally, that although I will occasionally make some claims about “what is asserted” in some case, I mean such claims only to have to do with my notion of asserted content. That is to say, when I make a claim about “what is asserted,” I mean to claim something relevant to whether a

---

24 The picture of assertion outlined here is heavily based on Stalnaker’s work, especially his (1978).

25 Stalnaker (2002) argues that there are some cases where what is in the common ground is determined not by what the conversational participants believe, but by some weaker belief-like attitude: what the participants accept, where “accept” denotes a family of attitudes including believing, presuming, assuming, and accepting for the purposes of an argument or an inquiry (Stalnaker 2002, 716). In such a conversation, the sincerity principle in the text will have to be amended: one asserts sincerely iff one accepts what one asserts. In general, the idea is that an assertion is the expression of some belief-like state; the assertion is sincere if one is actually in the relevant state. For simplicity, in the main text, I consider only cases where the relevant state is belief.
given assertion should be counted as sincere or not. It is very likely that there are other reasons why one might care "what is asserted" in a particular case which would give a different verdict; I hope to remain neutral on such questions.

2.4.1 Sincerity

A starting observation about sincerity: the sincerity of an assertion should depend on what the speaker intends the assertion to do, or what effect the speaker believes the assertion will have if accepted, rather than on what effect the assertion actually would have if accepted. The two come apart in cases where the speaker is mistaken about what the common ground is.

If an assertion of a sentence expressing the proposition \( p \) is accepted by its audience, the actual effect on the common ground will be to add \( p \). But what the speaker intends assertion of a sentence expressing \( p \) to accomplish is to update the context set (the set of worlds compatible with the common ground) to the intersection of the set of \( p \)-worlds with what the speaker thought the original context set was. If the speaker was mistaken about what the original context set was, this will not be the same as the actual effect of successful assertion of a sentence expressing \( p \).

Here are a pair of examples to illustrate. Rush and Dubya work in the same office. Dubya is a bit clueless, and Rush is self-absorbed and not very nice. Today is Friday.

Case 1: Rush needs to deposit his paycheque before the start of business on Monday. He has written a big cheque which will be cashed on Monday, and he must make sure there is enough money in his account to cover it. It would be hard for him to make it to the bank after work before the bank closes, but he has a lot of work to do and wants to avoid leaving early if possible.

All day, Rush has been complaining about his problem to his

\(^{26}\)By “common ground” I mean, approximately, what the speaker and hearer commonly believe: what they each believe, and believe the other believes, and believe the other believes they believe, and so on. For more on common ground, see Stalnaker (1978, 2002).
co-workers, and asking if anyone knows if the bank is open on Saturdays. He knows that as of two weeks ago, the bank was open on a Saturday, but he still worries that they might have changed their hours since.

Towards the end of the day, Rush runs into Dubya. Presuming that Dubya knows all about Rush’s problem—surely everyone knows by now—Rush asks if Dubya knows if the bank will be open tomorrow. But Dubya hasn’t heard anything about Rush’s predicament, and presumes that Rush just doesn’t know what the bank’s normal hours are. Dubya was at the bank on a Saturday two weeks ago, so he says, “Yes, they’re open Saturdays.” This, I take it, is a sincere assertion.

The actual effect of Dubya’s assertion is to bring Rush to believe that the bank has not deviated from its regular hours; the intended effect of Dubya’s assertion is to bring Rush to believe that the bank’s regular hours include Saturdays. We get the right verdict if we say that Dubya’s assertion is sincere iff he believes what he intends to bring Rush to believe by his assertion. We get the wrong verdict if we say that Dubya’s assertion is sincere iff he believes what he actually brings Rush to believe.

Case 2: Dubya needs to deposit his paycheque before the start of business on Monday [for the same reasons Rush did in Case 1]. It only occurs to Dubya at the end of the workday that some banks are closed on Saturdays, and he has no idea whether his bank is one of them. Rush knows that Dubya needs to deposit his cheque before the start of business on Monday, and Rush thinks that everyone knows that the bank was open on Saturdays as of two weeks ago. Therefore, when Dubya asks Rush if the bank will be open tomorrow, Rush presumes that Dubya is asking whether the bank has changed its hours. Rush is not a nice person, and wants to mislead Dubya, so he says “Yes, they’re open Saturdays,” despite not having any reason to think that the bank still has the same hours it had two weeks
ago. In fact, Rush believes that the bank has changed its hours, and will be closed tomorrow. I take it this is clearly an insincere utterance. But in fact what Dubya was asking was not whether the bank has stopped being open on Saturdays: it would suffice to answer poor clueless Dubya’s question to say that the bank was open on Saturday two weeks ago. Dubya is not considering the possibility that the bank’s hours have changed so recently.

The actual effect of Rush’s assertion is to bring Dubya to believe that the bank’s regular hours include Saturdays; its intended effect was to bring Dubya to believe that the bank still holds its regular hours (which include Saturdays). Rush does believe that the bank’s regular hours include Saturdays; he does not believe that the bank still holds its regular hours. We get the right verdict in this case if we say that Rush’s assertion is sincere iff he believes what he intends to bring Dubya to believe by his assertion. We get the wrong verdict if we say that Rush’s assertion is sincere iff he believes what he actually brings Dubya to believe.

So let us take it for granted that whether an assertion counts as sincere depends on what the assertor intends her assertion to do, rather than what it actually does. Now we turn to the context-dependence of what an assertion that \( p \) does. It is widely agreed that what an assertion aims to do is add a proposition to the common ground. But what this amounts to depends on what the common ground was before the assertion.

To see this, here is a claim about assertion: because speakers are generally aware of the possibility that they are mistaken about what the common ground is, what one aims to do in asserting that \( p \) is not just to add \( p \) to the common ground; rather, one aims to update the context set to (a subset of) the intersection of the set of \( p \)-worlds with what one takes the context set to be. Some examples, with pictures, will help illustrate the point.
2.4.2 Expressed and Asserted Content

Before we get to the pictures, here is an overview of the picture of assertion I favour. A sentence, uttered in a certain way in a certain context, has a certain *expressed content*—generally, a proposition. This is pretty vague, and intentionally so: I have no insight to offer about which propositions are expressed by which utterances of which sentences; I hope to remain neutral in such debates. For any assertive utterance, pick your favourite story about what proposition was expressed, and that will work for my purposes.  

In general, expressed content need not be the same as semantic value—in particular, it will be different in cases of idiom, sarcasm, and perhaps some cases of implicature. An utterance of, “Right, Fred’s a good dean and I’m a monkey’s uncle,” intuitively expresses the proposition that Fred is not a good dean. There are complicated things determining the expressed content of any given utterance, but if you want to know much more about it, you’ll have to find another philosopher, or a linguist.

Here is where I have a story to tell: a proposition, expressed assertively in a certain context, has a certain *asserted content*. There are thus two stages in getting from an utterance to what that utterance asserts; there are contextual factors involved in both stages (in the first stage, this will include, e.g., determining the referents of indexicals), but the kind of context-sensitivity that interests us here comes in the second stage. A proposition

---

There might be some difficulty if you happen to be Herman Cappelen or Ernie Lepore, however. According to Cappelen and Lepore’s Speech Act Pluralism, there is no such thing as the proposition expressed by a given utterance; and furthermore, there is no algorithm that will give all the propositions expressed (said, asserted, etc.) by an utterance. However, at least as of 2005, Cappelen and Lepore suspect that no one else is a Speech Act Pluralist of this sort, so I shall try not to be too worried (Cappelen and Lepore 2005, 5). Furthermore, Cappelen and Lepore (2005, 202) explicitly reject the sincerity principles I build my argument on: “Speakers need not believe everything they sincerely say.” So Cappelen and Lepore got off the bus well before this note.

I do think that some version of my argument might work even for Speech Act Pluralists, provided we appropriately restrict our sincerity principles—though this would of course be difficult to manage. Roughly, the idea is that if, for a given assertive utterance, we (somehow) pick an appropriate member of the set of propositions expressed by that utterance, the sincerity principles I rely on will hold for that proposition; furthermore, if we plug that proposition in as expressed content in my sense, we can run the argument to follow on the asserted content generated thereby.
can in general be identified with a set of possible worlds;\(^{28}\) the intersection of that set with some other contextually-determined set of possible worlds (e.g., the set of worlds compatible with the common ground as the speaker takes it to be) determines the asserted content. Now, on to the pictures.

![Diagram of possible worlds](image)

(a) \(\neg p \ p\)

(b) \(\neg p \ p\)

Figure 2.1

Let \(U\) (the big box) be the universe of possible worlds. Figure 2.1a shows the proposition expressed by \(p\): the area of the box not shaded out represents the worlds where \(p\) is true. This is what I call the expressed content.

In Figure 2.1b, the oval (labelled “C”) represents the worlds compatible with the common ground. The area outside the oval is shaded in darkly; the area inside the oval on the not-\(p\) side of the universe is shaded in more lightly. The idea is that the darkly shaded area represents worlds that are considered to have been ruled out even before the speaker’s assertion; then the assertion itself aims to update the common ground to include only the unshaded area, which is inside the oval and also on the \(p\) side of the universe. This represents what I call the asserted content.

Now we are in a position to show why we should not take the aim of an assertion that \(p\) to be simply adding \(p\) to the common ground. If that were the case, we would take the asserted content to be the entire section of the universe to the left of the vertical line, because, again, the speaker may be mistaken about what the common ground is. Suppose, as in one of Schaffer’s examples, our conversational participants are talking about what kind of bird

\(^{28}\)This identification involves a certain simplification or idealization; see §2.5.1 for more on this.
is in the garden. Consider Figure 2.2a. Here, \( p \) is the proposition that the bird is a canary; the circle labelled “\( T \)” represents worlds where there is a canary in the garden—specifically, Tweety Bird, a cartoon canary. Suppose the speaker (mistakenly) thinks the common ground rules out everything outside the oval labelled “\( C \)”, including the circle labelled “\( T \)”. In that case, aiming to update the common ground to allow only worlds where there is a three-dimensional canary in the garden, the speaker will utter a sentence like “It’s a canary,” rather than a sentence like “It’s a three-dimensional canary.” Speakers count on the common ground to do some work for them. Now, if the result of uttering “It’s a canary” is the situation depicted in Figure 2.2a, the assertion was not successful; the speaker has not achieved her aim. The common ground has not been updated as desired. Rather, a successful assertion would result in the situation of Figure 2.2b, with the Tweety-worlds ruled out as well.

![Figure 2.2](image)

So, again, we should identify the asserted content with the set of worlds compatible with the common ground as the speaker takes it to be where the expressed proposition is true.

So much for asserted content. What is the role of expressed content in this picture? For our purposes, what matters about expressed content is the following: we say that one has asserted that \( p \) just in case one has made an assertive utterance with the expressed content \( p \). That is, to say that one has asserted that \( p \) is to specify the expressed content of one’s assertion,
not its asserted content. The notion of asserted content is designed to get at what one intends to do by asserting, with a view to getting at what one must believe in order for one’s assertion to be sincere; but our normal way of talking about what people assert does not have this specialized aim. In the Tweety case, it is more natural to say that by uttering “The bird is a canary,” I have asserted that the bird is a canary (p in figure 2.2b) than that I have asserted that the bird is a three-dimensional canary (the proposition corresponding to the unshaded area in figure 2.2b).

A more compelling case for using expressed rather than asserted content to determine when one has asserted that p comes from considering pairs of cases where, intuitively, one asserts the same thing despite the common ground being different (i.e., despite the asserted content being different). Consider the following variations on the bird-in-the-garden case.

I am on the telephone with you, describing what I am watching on television. It is a program you have never seen or heard of, and which you cannot watch where you are. There is a bird of some kind that appears on screen. I describe the bird, and ask you what kind it is.

**Variant 1**: You think I am watching a cartoon. The description fits Tweety bird, a cartoon canary. You say, “It’s a canary.”

**Variant 2**: You think I am watching a live-action show. The description fits a (three-dimensional, actual) canary. You say, “It’s a canary.”

**Variant 3**: You don’t know whether I am watching a cartoon or a live-action show, and don’t care enough to find out. The description fits a canary, or a cartoon canary. You say, “It’s a canary.”

It seems to me that, intuitively, in all three cases you have asserted (the proposition) that the bird on the television is a canary. That is, it is intuitive that there is a proposition p such that, in all three cases, you have asserted that p. However, you mean to eliminate different possibilities in each case, and to leave different possibilities uneliminated: in Variant 1, but not in...
Variant 2, you mean to rule out the possibility that the bird on the television is a cartoon duck; in Variant 2, but not in Variant 1, you mean to rule out the possibility that the bird on the television is a three-dimensional, actual duck; in Variant 3, you mean to eliminate both possibilities. Thus, the assertions in the three cases have different asserted content. Therefore, we get the wrong result if we say that one asserts that $p$ just in case one utters assertively something with the asserted content $p$. On the other hand, if we say that an assertion that $p$ is one with the expressed content $p$, we do not run into this problem, and we also match our ordinary practice of describing assertions.

### 2.4.3 The Argument

Now we have all the pieces we need to get to the context-sensitivity of belief. An assertion that $p$ is sincere iff the assertor believes that $p$; equivalently, an assertion is sincere iff the assertor believes what she asserts. But what one asserts when one asserts that $p$ depends on context. In particular, it depends on what the assertor takes the common ground to be. An assertion that $p$ has different content in different contexts; therefore, the belief an assertor must have for an assertion that $p$ to count as sincere will be different in different contexts. But what one must believe in order that an assertion that $p$ count as sincere is that $p$; therefore, what belief that $p$ amounts to depends on context. To believe that $p$ is to believe that the actual world is a member of the subset of the context set where $p$ is true. This is just the sort of (believer-centred) “contextualism” about belief we set out to establish, i.e., sensitivism.

Here is the argument again, more carefully. Suppose $S$ has asserted that $p$. This means that $S$ has made an assertive utterance with the expressed content $p$. $S$’s assertion is sincere if and only if $S$ believes that $p$. But whether $S$’s assertion is sincere or not depends on $S$’s communica-

---

29I say these two versions of the sincerity principle are equivalent; what I mean is that they are *intuitively* equivalent. That is, I cannot imagine finding one principle intuitively appealing without finding both principles intuitively appealing. However, I do not mean to claim that the best precisifications of each principle are logically equivalent.
tive intentions—what effect $S$ intends her assertion to have on the common
ground of the conversation—and this is not reflected in the expressed content
of her utterance. To capture $S$’s communicative intentions, we must look at
the asserted content of her utterance, which depends on what $S$ thinks the
common ground is. This, the asserted content of $S$’s utterance, tells us what
belief $S$ must have for her assertion to be sincere. Therefore, what belief $S$
must have for her assertion to count as sincere depends on what $S$ thinks the
common ground is. But we have already said that $S$’s assertion is sincere if
and only if $S$ believes that $p$. Therefore, whether $S$ believes that $p$ or not
depends on what $S$ thinks the common ground is. This is the conclusion we
hoped for.

Our conclusion here establishes a form of context-sensitivity of belief,
but does it establish sensitivism? Recall that the sensitivist does not claim
merely that people have different beliefs in different contexts, but that whether
$S$ believes that $p$ depends, in part, on what alternative possibilities $S$ is tak-
ing seriously. The above argument concludes that whether $S$ believes that $p$
in a context depends, in part, on what $S$ takes the common ground to be in
that context. I think this does amount to a dependence on what alternatives
$S$ is taking seriously: arguably, what it is for a possibility to be unexcluded
by the common ground (i.e., for a possibility to be included in the context
set) is for the agents involved to be taking that possibility seriously. If $S$
thinks all parties to the conversation are taking some possibility seriously,
then she must think that she herself is taking that possibility seriously as
well. $S$ thinking that $S$ is taking a possibility seriously is not quite the same
thing as $S$ taking the possibility seriously, but it’s pretty darn close, and in
most cases will amount to the same thing. Perhaps there are cases where $S$
thinks that she is taking a possibility seriously, but actually is not; however,
I cannot think of such a case, nor can I convince myself that a serious objec-
tion to the present view based thereon can be offered. If such an objection
is possible, I will leave it to other philosophers to find it.

The present argument for sensitivism is not susceptible to the objections
of §2.3. There is no dependence on the knowledge account of assertion,
the variation of assertability across contexts need not be due to shifting
epistemic standards, and the inference involved is not an inference to the
best explanation. To reject the present argument, one must reject one of our
principles about assertion and sincerity, or reject our picture of assertion.
To be sure, that picture of assertion is largely based on inference to the best
explanation—but the data to be explained here are very general pragmatic
phenomena, which have been well studied for some time now. There is
wide consensus that the Stalnakerian common-ground picture of assertion is
correct. This is in contrast to the data to be explained in the case of DeRose
and Schaffer’s arguments for contextualism about knowledge. There, the
data to be explained are themselves controversial: it is the alleged knowledge
norm on assertion that is supposed to be best explained by contextualism
about knowledge.

In §2.2, I claimed that sensitivism about belief is at least as well sup-
ported as contextualism about knowledge by DeRose/Schaffer-style consid-
erations about assertion. Now we see that sensitivism about belief is in fact
better supported than contextualism about knowledge: as well as modifica-
tions of the DeRose/Schaffer arguments, which should be acceptable if the
original arguments were acceptable, we also have an alternative argument
which should be acceptable even to those who reject the knowledge account
of assertion or think the knowledge norm can be explained without appeal
to contextualism. Furthermore, the prospects for an analogue of the present
argument with knowledge substituted for belief appear dim: there is no tie
between knowledge and assertion as close as the tie between belief and sin-
cere assertion to which we appeal here. The sensitivist is strictly better off
here than the knowledge-contextualist.

2.5 A Model for Belief

I have presented the foregoing as an argument for sensitivism, but we can
draw a stronger conclusion. That is, the considerations of the previous sec-
tion do not push us to accept just any sensitivist model of belief. In this
section, I offer a way of modelling belief states which easily makes sense of
the picture of sincere assertion given above.
2.5.1 Some Simplifications Eliminated

To begin, there are some simplifications I have made up to this point for reasons of clarity, which can now be dispensed with. In particular, I have assumed that successful assertions only increase the common ground, and reduce the context set. This is, of course, not the case. For example, consider again the case where the conversants are wondering what kind of bird is in the garden. An assertion that it might be a squirrel rather than a bird does not aim to eliminate any possibility; rather, it raises a new possibility previously excluded. Prior to the assertion, it was taken for granted by all parties that the creature in the garden is a bird; if the assertion is successful, then all parties will now suppose that it is either a bird or a squirrel. This means an enlargement of the context set, and a reduction in the common ground (as we have been understanding it), since squirrel-possibilities are now under consideration. And there are some assertions which aim only to point out or make salient that certain things are already in the common ground; assertions of the form “As we all know, ...” or “Of course, we’re taking for granted that ...” are often of this sort. (Cf. Abbott 2008.)

Another simplifying idealization is built into my use of possible worlds, and my representing propositions as sets of possible worlds. The trouble with this is that it requires us to represent logically equivalent propositions as identical. The same set of possible worlds corresponds to both $p$ and $(p \land q) \lor (p \land \neg q)$. Therefore, if we represent an agent as believing the former, we must also represent her as believing the latter, since we represent both by the same object. In other words, representing propositions as sets of possible worlds means building in logical omniscience.

The problem here bears some similarity to a family of problems of logical omniscience run into by formal epistemologists of various stripes—I have in mind Bayesians and epistemic logicians, in particular—but the difference between my project and theirs makes certain familiar moves unavailable to me. I am interested in understanding and representing belief in general, whether rational or not; but formal epistemologists are usually concerned with representing more ideal agents. This gives them some excuse for building logical
omniscience into their models: it is at least debatable whether ideally rational agents must possess perfect logical knowledge. But it is a clear problem if a model of non-ideal, not necessarily rational belief has it that all believers are logically omniscient.

However, the modesty (in a sense) of my goal also opens up a way of dealing with the problem of logical omniscience which is unavailable to Bayesians and epistemic logicians. Let me first explain how I deal with the problem, and then explain why this strategy is unavailable to those others.

We must modify the idea of possible worlds at work in our model. In fact, I will cease calling them possible worlds, since we will give up any connection with logical or metaphysical possibility; instead, let us speak of points in the model. We have a universe (set) of points, and every proposition has a truth value at each point. Thus, we can still represent propositions as sets of points.

The believer’s logical knowledge will determine the population of points. If the agent knows that \( p \) and \( (p \land q) \lor (p \land \neg q) \) are logically equivalent, then there will be no point where one is true and the other false; otherwise, there will be such points.\(^{30}\) Thus, the believer’s logical knowledge also determines the individuation of propositions: if the agent believes that \( p \) and \( q \) are equivalent, then both propositions will be represented by the same set of points; so, as far as the model is concerned, they are the same proposition. Once we have this clarification of what the points in our model are, everything works just as it would if we were using possible worlds.

The above strategy for dealing with the problem of logical omniscience is open to us because we are not giving a logic. If we were, we would need some recursive rule for determining how complex propositions are represented in the model. For example, typically, if \( p \) and \( q \) are represented, respectively, by sets of points \( P \) and \( Q \), then \( p \land q \) would be represented by the intersection of \( P \) and \( Q \). Thus, we would not have the option of representing \( p \land q \) by anything but the set of worlds where both \( p \) and \( q \) are true; all agents would have to recognize conjunction introduction and elimination as valid.

\(^{30}\)It would probably be better to say that the believer’s logical beliefs determine the population of points, but that just sounds confusing.
so to speak. But I do not have any need to construct complex propositions from simple ones; for my purposes, there is no problem with regarding all propositions as simple. Therefore, any proposition can be represented by an arbitrary set of points. The set of points where \( p \land q \) is true need not be the set of points where both \( p \) and \( q \) are true.

To be sure, representing propositions as sets of points in the model still runs into some problems analogous to those that arise in using sets of possible worlds. In particular, any two propositions which the believer’s logical beliefs treat as necessarily true will be represented as identical. This may make the model unsuitable for representing, say, beliefs about mathematics. I will not offer a solution of this problem; I concede that the way of modelling belief on offer here is only useful for representing a restricted range of beliefs. However, I think the restriction is not so severe as to remove all usefulness or interest from the models—we can, after all, still represent beliefs about brains in vats, bank hours, and indeed anything else in the physical world.

2.5.2 The Model

Now we are ready to present the model.

Definitions

Let a doxastic state \( S \) for a set of propositions \( P \) be a quadruple \((U_S, C_S, \| \cdot \|_S, \preceq_S)\), where \( U_S \) is a set of points, \( C_S \) is a set of non-empty subsets of \( U_S \) (“contexts”), \( \| \cdot \|_S \) is a valuation function taking members of \( P \) to subsets of \( U_S \), and \( \preceq_S \) is a reflexive,\(^{31}\) transitive,\(^{32}\) and \( C_S \)-well-founded\(^{33}\) relation among the members of \( U_S \).\(^{34}\) “\( x \preceq y \)” is meant to be read as, “\( y \) is ruled out at least as strongly as \( x \) (according to the agent whose beliefs are being represented).” It might also be helpful—though not quite accurate—to think of “\( x \preceq y \)” as saying that \( x \) is more plausible than \( y \) according to the agent; I

---

\(^{31}\)For any point \( x \), \( x \preceq x \).

\(^{32}\)For any points \( x, y, z \), if \( x \preceq y \) and \( y \preceq z \), then \( x \preceq z \).

\(^{33}\)For any \( C \in C_S \), there is a minimal element \( m \in C \): for all \( x \in C \), \( m \preceq x \).

\(^{34}\)The subscripts on \( U, C, \| \cdot \|, \) and \( \preceq \) will be suppressed where there is no ambiguity which doxastic state is in question.
will say more about this shortly. For a proposition \( p \in P \), \( \parallel p \parallel \) is the set of points in \( U \) at which \( p \) is true. If \( x \) is a point with \( x \in \parallel p \parallel \), then we will write \( x \models p \); if \( A \) is a set of points with \( A \subseteq \parallel p \parallel \), then we will write \( A \models p \).

Let a context be represented by a context set \( C \in \mathcal{C} \). Because \( \preceq \) is well-founded, we can always find, for any \( C \), a minimally ruled-out (maximally plausible) subset \( B_C \subseteq C \); that is, there is a \( B_C \subseteq C \) whose members are those \( x \in C \) such that for all \( y \in C \), \( x \preceq y \). Then we say that \( p \) is believed in context \( C \) if and only if \( B_C \models p \). That is, \( p \) is believed in \( C \) just in case it is true at all the least ruled-out (most plausible) \( C \)-points.

**Discussion**

The model defined above is meant to be a generalization of the picture we saw in §2.4 to cases not involving assertion. The context sets \( C \in \mathcal{C} \) in the model obviously correspond to the Stalnakerian context set generated by the information in the common ground of a conversation: just as the common ground restricts interlocutors’ attention to a subset of the space of possible worlds, the context sets \( C \in \mathcal{C} \) in our model restrict a believer’s attention to a subset of the points populating her doxastic state \( \mathcal{S} \). Not all subsets of \( U \) clearly correspond to any recognizable context (this is why \( \mathcal{S} \) includes the distinguished set of contexts \( \mathcal{C} \)). For instance, a singleton set containing only a single point, with nothing ruled out (or in) is hard to understand as corresponding to anyone’s beliefs in any real life context. It seems to me that genuine belief always involves ruling something out; but even if this is not the case, our framework can handle it—if every subset of \( U \) corresponds to some possible context, then let \( \mathcal{C} \) be the power set of \( U \), the set of all subsets of \( U \). The important thing is that we have the option of excluding some subsets of \( U \) as not corresponding to any real context.

The relation \( \preceq \) can be seen as generalizing the shading used in §2.4. Figures 2.1–2.2 used at most three colours: dark grey for possibilities\(^{35}\) ruled out by the common ground, light grey for possibilities ruled out by the assertion of \( p \), white for possibilities not ruled out. Context-independent

---

\(^{35}\)I promised to stop referring to points in the model as “possible worlds,” but “possibilities” is just too convenient a term to give up completely.
doxastic states, on this model, can be visualized as using arbitrarily many colours: if $S$ has it that $x \preceq y$, then shade $y$ darker than $x$. In a context $C$, these shadings on the doxastic state can be used to generate a model in the style of §2.4: everything outside of $C$ is shaded dark grey; the lightest possibilities within $C$ (i.e., those in $B_C$) are shaded white; and everything else in $C$ is shaded light grey. Of course, dark grey, light grey, and white will not mean quite the same thing in this model as in that of §2.4, since we are now working with a model for belief, not assertion. Here, then, dark grey would indicate points unconsidered, or not taken seriously in the context; light grey would indicate points considered but ruled out (disbelieved); and white would indicate points considered, and not ruled out (believed).

This explains why we required $\preceq$ to be reflexive, transitive, and $C$-well-founded. Reflexivity and transitivity make $\preceq$ nicely behaved as a (non-strict) ordering: there are no loops with $x \prec y$, $y \prec z$, and $z \prec x$.36 $C$-well-foundedness ensures that any context $C \in \mathcal{C}$ will have a non-empty minimally ruled-out subset $B_C$. Now, it might seem that requiring transitivity is not just an innocent formal convenience. That is, someone might think that $x$ is more likely than $y$ in a context $C_1$ where $z$ is not under consideration, and likewise think that $y$ is more likely than $z$ if $x$ is not under consideration, and that $z$ is more likely than $x$ if $y$ is not under consideration (contexts $C_2$ and $C_3$, respectively); but it seems that to represent such an agent’s doxastic state, we must violate transitivity. But this is not the case, for we can always introduce a fourth point, $z'$, which agrees with $z$ on every proposition $p \in P$, but with $z' \prec x \prec y \prec z$ (see Figure 2.3, where an arrow from a point $v$ to a point $w$ indicates that $w \prec v$).37 Then, if we take $C_1 = \{x, y\}$, $C_2 = \{y, z\}$, and $C_3 = \{x, z'\}$, we get a representation of the agent as described, but without having to violate transitivity. This is possible because there is no

---

36 Here, of course, “$a \prec b$” means $a \preceq b$ and $b \notin a$.

37 Incidentally, note that so far, I have tacitly assumed that there is no context in which the agent takes all three possibilities $x, y$ and $z$ into consideration; if there is such a context, then whichever possibility is (or possibilities are) not ruled out will be the one that the new point (or points) agrees with on all propositions. When there is no such synoptic context, the three original possibilities are in a symmetrical situation, and it makes no difference which one is “copied” by the new point(s).
requirement that distinct points of \( U \) must differ in the propositions they make true.

Figure 2.3

Here is another way of putting the point. A doxastic state \( S \) for the propositions \( P \) is a good model of an agent \( S \) if and only if \( S \) gets \( S \)'s pattern of belief and non-belief in the propositions \( P \) right across the contexts represented by the members of \( C_S \). The points \( x \in U_S \) are useful for representing those propositions and contexts, but do not have any independent content; outside the context of the model \( S \), it does not make sense to ask about whether the agent \( S \) believes, say, that \( x \) is actual. Thus, without violating transitivity, we can build a doxastic state which accurately models an agent who believes \( p \) rather than \( q \) in a context where \( r \) is not a serious alternative, believes \( q \) rather than \( r \) when \( p \) is not a serious alternative, and believes \( r \) rather than \( p \) in a context where \( q \) is not a serious alternative. Just take the doxastic state of Figure 2.3b, with \( \| p \| = \{ x \}, \| q \| = \{ y \}, \) and \( \| r \| = \{ z, z' \}. \) Then the context \( C_1 = \{ x, y \} \) will be one where the agent believes \( p \) rather than \( q \), \( C_2 = \{ y, z \} \) one where the agent believes \( q \) rather than \( r \), and \( C_3 = \{ x, z' \} \) one where the agent believes \( r \) rather than \( p \). That is to say that in \( C_1 \), the agent believes \( p \) and rules out the only serious \( q \)-possibility, and takes for granted that \( r \) is false, or in other words, does not take any \( r \)-possibility seriously. Thus, the transitivity constraint does not
reduce the expressive power of doxastic states.

Note that the different levels or shadings of the points in $U$ induced by $\preceq$ should not be thought of as degrees of belief, at least in the usual sense. Rather, they should be thought of as dispositions to believe. What it means for some point $x$ to be shaded more lightly than some other point $y$ is not that the agent believes $x$ is the case more strongly or with more certainty than $y$; rather, what it means is that in a context where attention is restricted to the points $x$ and $y$, the agent will believe those things true in $x$, regardless of what is true in $y$. This is part of why I said above that it is better to think of “$x \preceq y$” as saying that $x$ is less ruled out than $y$, rather than as saying that $x$ is more plausible than $y$ (according to the agent). To illustrate, here is a plausible (but highly simplified) way of modelling many people’s attitudes to skeptical scenarios; I have in mind people who feel the force of skeptical arguments.

Let $U = \{2H, 1H, 2H_{BIV}, 1H_{BIV}\}$. These points correspond, respectively, to: the world where I have two hands and nothing else is unusual ($2H$); the world where I have only one hand, having lost one in an accident ($1H$); the world where I am a handless brain in a vat having experiences as of having two hands ($2H_{BIV}$); and the world where I am a handless brain in a vat having experiences as of having lost one hand in an accident ($1H_{BIV}$).

The worlds $2H$ and $2H_{BIV}$ are indistinguishable to me, as are $1H$ and $1H_{BIV}$: for each pair, I have exactly the same experiences in both members of the pair. For convenience, let us write “$x \simeq y$” when $x \preceq y$ and $y \preceq x$. Thus, we will have $2H \simeq 2H_{BIV}$ and $1H \simeq 1H_{BIV}$. Furthermore, let us suppose that I do seem to have two hands—that is, I have an experience as of having two hands. I see two hands before me, I have no memory of losing a hand in an accident, and so on. Then I will be inclined to rule out any point where I seem to have one hand rather than any world where I seem to have two hands: we will have $2H \preceq 1H, 2H_{BIV} \preceq 1H_{BIV}, 2H \preceq 1H_{BIV}$, and $2H_{BIV} \preceq 1H$. This gives us the situation of Figure 2.4.

Now, in most ordinary contexts, the skeptical possibilities $2H_{BIV}$ and $1H_{BIV}$ will not arise. Thus, an ordinary context set $C_{\text{ordinary}}$ will only include $2H$ and $1H$. Since $2H$ is the unique minimal element of such
a $C_{\text{ordinary}}$, we will have $B_{C_{\text{ordinary}}} = \{2H\}$; so the agent will, in ordinary contexts, believe whatever is true in $2H$—in particular, that she has two hands. This is depicted in Figure 2.5a. However, by raising skeptical possibilities to salience, an interlocutor can expand the context set to a new $C_{\text{skeptical}}$ which includes $2H_{BIV}$. Since $2H_{BIV} \simeq 2H$, we will have $B_{C_{\text{skeptical}}} = \{2H, 2H_{BIV}\}$, so that the agent only believes whatever is true in both $2H$ and $2H_{BIV}$—in particular, she will not believe that she has two hands, since this is false at $2H_{BIV}$, though neither will she believe that she has no hands, since this is false at $2H$. This is the situation depicted in Figure 2.5b.\textsuperscript{38}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{context.png}
\caption{(a) Ordinary and (b) skeptical contexts.}
\end{figure}

What I mean when I say that the ordering given by $\preceq$ does not give an ordering of degrees of belief should now be clear. The effect of skeptical conversational manoeuvre like saying “But how do you know you’re not a brain in a vat, stimulated to have experiences as of having hands?” is not to bring the skeptical brain-in-a-vat possibility to equal plausibility as the

\textsuperscript{38}In Figure 2.5b, $1H_{BIV} \in C_{\text{skeptical}}$. This is somewhat arbitrary; it is not clear to me whether, in general, skeptical arguments bring to salience skeptical versions of possibilities already ruled out. In any case, nothing here turns on the inclusion or exclusion of $1H_{BIV}$ in the skeptical context set.
ordinary alternative; rather, it is to destroy belief in the ordinary alternative by bringing into consideration an alternative possibility one cannot rule out.

Before we move on to generalize from the particular skeptical case described above, a word about ruling out is in order. One often hears about ruling out possibilities from proponents of relevant alternatives (RA) approaches to knowledge. The sort of ruling out invoked to explain knowledge is different from the sort involved here. We might say that RA theorists are primarily concerned with an epistemic sort of ruling-out, whereas I am primarily concerned with a psychological sort of ruling-out. Thus, an RA theorist of a certain (contextualist) stripe might diagnose skeptical arguments as being effective because they bring to relevance remote possibilities about which one has no evidence, and so which one cannot rule out; since one cannot rule out a relevant alternative to one’s having hands, one cannot know that one has hands. On the other hand, it certainly is possible to rule out in my sense the possibility that one is a brain in a vat. I suspect the skeptic would have little success destroying, say, my grandmother’s belief that she has hands (to say nothing of what the skeptic might do to my grandmother’s knowledge that she has hands); it is a fact about her psychology that it would be very difficult to make her take seriously the possibility that she is a brain in a vat. Thus, the model of Figure 2.5b simply would not describe her, even when the skeptic gets to work on her. On the other hand, a certain proportion of undergraduates in introductory epistemology classes are very easy to put in the situation described by Figure 2.5b. The short of it is: my models say nothing about one’s reasons for ruling out a possibility, or for ignoring it.

Now, back to our example. It is, of course, highly simplified. Here are the important parts, though, for our diagnosis of effective skeptical arguments: we have a population of ordinary worlds (1H and 2H); a population of skeptical worlds where a target proposition (that I have hands) is false (1H_BIV and 2H_BIV); the skeptical worlds have ordinary counterparts from which they are indistinguishable (1H_BIV ≃ 1H and 2H_BIV ≃ 2H); ordinary context sets include only ordinary worlds, but bringing skeptical possibilities to salience can expand the context set to include some skeptical worlds;
in particular, some skeptical worlds will be included in the new belief set, $B_{C,\text{skeptical}}$. This pattern will generalize. Effective skeptical arguments are generally effective precisely because they point to worlds which are indistinguishable from those we would otherwise believe are actual; we have no reason to rule out the skeptical worlds without also ruling out the ordinary worlds, and so we wind up accepting the skeptical possibilities as legitimate and un-rulled-out. Thus, belief in a target proposition is destroyed.

But this does not require that we have equal confidence in the skeptical and ordinary possibilities. It only requires that we do not rule out the skeptical possibilities. This is why the ordering given by $\preceq$ is not an ordering on degrees of belief, but rather indicates conditional dispositions to believe. We will see a model for degrees of belief in chapter 4.

The preceding example is meant primarily to serve as an illustration of how my doxastic states are supposed to work, but it also has some intrinsic interest of its own. In particular, it is supposed to be an advantage of contextualism that it explains the effectiveness of skeptical arguments. (Briefly: the skeptic puts her victim into a context in which skeptical scenarios are salient, and this raises the standards for knowledge; in such a context, ordinary people do not know even ordinary facts about their own hands; the skeptic then leads us to conclude, illegitimately, that we never know ordinary facts about our own hands; but on the contrary, in ordinary contexts, the standards for knowledge are not so high, and ordinary people do, after all, ordinarily know ordinary things about their hands.) The example we have just worked through gives a similarly appealing diagnosis of the effectiveness of skeptical arguments without committing us to any controversial theory of knowledge or of “know”. Just as the contextualists say, skeptical arguments serve to make salient remote possibilities, e.g., that one is a brain in a vat; this destroys belief in ordinary alternatives to those possibilities, e.g., that one has hands. Lacking belief, one also lacks knowledge, and the skeptic has succeeded.

However—and this may be a strength, as well—unlike the contextualist response to skepticism, the present explanation says nothing about why or whether skeptical arguments are bad. That is, what I offer here is not so
much a response to skepticism as a diagnosis of its effectiveness. I say this
may be a strength because there have been a number of objections posed in
the literature to the contextualist response to skepticism (e.g., Bach 2005,
Feldman 2001, 2004, Klein 2000); it seems to me that what is appealing
about the contextualist response is the diagnostic part—it makes sense to
think that skeptics undermine our knowledge of ordinary things by bringing
to light remote possibilities.\footnote{This should not be confused with the claim that the contextualist response correctly
diagnoses the disagreement between the skeptic and the dogmatist; on the contrary, the
contextualist response is notably criticized for getting the latter diagnosis wrong in Feld-
man (2004) and Klein (2000). Rather, the claim is that the contextualist response correctly
diagnoses the skeptic’s effect on her victim (i.e., on someone who is persuaded by, or at
least feels the force of, the skeptical argument).} But this diagnosis is compatible with multiple
senses of “undermining”: maybe the skeptic reveals to us that we never know
very much, or maybe the skeptic tricks us into thinking so by causing us to
lose knowledge temporarily through loss of belief.

Finally, now that we have our model, we can give a precise explication
of the sincerity principle appealed to in §2.4. That principle was: An as-
sertion that $p$ is sincere if and only if the assertor believes that $p$. Let
$S = (U, C, \| \cdot \|, \preceq)$ be the doxastic state of an agent who has just made an
assertive utterance with the expressed content $p$. Then, the agent has as-
serted sincerely that $p$ if and only if $B_C \models p$, where $C \in C$ is the context set
corresponding to what the agent thinks is the common ground (i.e., $C$ is the set of all points where the information the agent thinks is in the common
ground is true). Recall that “$B_C \models p$” is to be read as “[the agent] believes
that $p$,” and it is easy to see that this is a formalization of the principle in
question. Furthermore, note that since $C$ is determined by what the assertor
thinks the common ground is, there is a clear sense in which the condition
$B_C \models p$ means that the assertor’s beliefs mirror the asserted content of her
utterance. After all, the asserted content corresponds, roughly, to $C \cap \| p \|$;
so we might say that our formalized principle requires that the agent rule
out all points outside the asserted content—i.e., that the assertor’s beliefs
rule out all possibilities she intends her assertion to rule out.
Assessment

Before we move on, let us assess the strengths and limitations of the model offered here. As I just mentioned, one limitation is that it is entirely qualitative, in the sense that it does not deal with degrees of belief at all. This limitation will be addressed in chapter 4. Furthermore, the model is entirely static: it is designed to represent an agent’s beliefs at a particular time, but has nothing to say about how an agent’s beliefs might change over time. This limitation I will not address in the present work, because it derives from a third limitation. This is a peculiar quirk of my project: I aim to give an account of belief which will be of use and interest to epistemologists, but I am primarily concerned with the nature of belief, not just rational belief. Without the restriction to rational belief, it is hard to put constraints on how a belief state might evolve over time; hence, my model is merely static.

But one reason why I hope this model will nevertheless be of interest to epistemologists is that it provides a framework for an account of rational belief. We might require, say, that certain beliefs be held in all contexts, perhaps via some requirements on the makeup of the universe $U$. Or we might put some additional constraints on $\preceq$. Or we might have some requirement on how one responds to context-shifting manoeuvres, or to new evidence. I hope to have gotten clearer on what belief is like partly as a means to getting clearer on what rational belief is like.

The model has other, more specific strengths too: it is designed to allow an intuitive treatment of sincere assertion, of course; and I have argued that it allows a nice explanation of what skeptical arguments can do. Furthermore, as I shall argue in the next chapter, it provides a nice treatment of the preface paradox.
3.1 Introduction

In this chapter, I use sensitivism to give a solution to the preface paradox (Makinson 1965), and compare my solution with others. In brief, according to my preferred solution, there is a change of context between the preface disclaimer and the body of the text; using the account of sincere assertion put forward in chapter 2, I argue that it follows that the preface writer does not express contradictory beliefs. But I am getting ahead of myself; this synopsis will make no sense without a statement of the paradox. So, to begin, here is a statement of the paradox:

Jennifer has written a history book. The body of the book consists of a large number of sincerely-made claims about history, each of which she has carefully researched. Each claim in the body of the book is one which Jennifer believes. However, Jennifer also writes a preface to the book in which, after thanking various colleagues for their help, she claims—again, sincerely—that she is bound to have made mistakes somewhere in the body of the book, and excuses her colleagues for them. That is, she believes that some of her claims in the body of the book are false.

It seems that Jennifer must have inconsistent beliefs. She believes each of the claims in the body of the book, but she also believes that at least one of those claims is false.

Jennifer’s situation constitutes a paradox because it is mundane—modest prefaces are hardly unusual—but describes a violation of what is widely regarded as a minimal constraint on rational belief: “don’t be inconsistent.”
I say this is a *minimal* constraint on rational belief because obeying it does not mean that one’s beliefs are particularly likely to be true; but disobeying it means that it is impossible for all of one’s beliefs to be true.\textsuperscript{40,41}

The preface paradox also involves a violation of a stronger constraint: “believe all the deductive consequences of your beliefs.” This constraint, often referred to as a “closure”\textsuperscript{42} constraint since it requires that one’s set of beliefs be closed under deductive consequence, is usually offered only as a constraint on ideal rationality, since it would be unrealistic to expect actual human beings even to be aware of what all the deductive consequences of their beliefs are, much less believe them all. But as an ideal, closure is not so implausible: it only requires that one believe things to which one is, in a sense, already committed. That is, if closure requires that one add some proposition $p$ to one’s set of beliefs, then $p$ must be true in any world where all of one’s previous beliefs are true. Since there is no way for $p$ to be false while the rest of one’s beliefs are true, adding $p$ to one’s set of beliefs does not introduce any new risk of error. Thus, by following closure, one has the opportunity to be right about more things without risking being wrong about anything one wasn’t already wrong about. The violation of closure in

\textsuperscript{40}To be sure, consistency is still a condition on *ideal* rationality, since it would take tremendous computational power to determine whether even a relatively small set of propositions is consistent—power that actual human beings cannot hope to possess (Cherniak 1986, 93). Thus, the sense of minimality in which consistency is a minimal constraint on rationality is not the same as the sense of minimality that figures in Cherniak’s *Minimal Rationality* (Cherniak 1986); the latter sense of minimality is opposed to ideality, and has to do with taking human limits seriously. For other works which approach rationality in the same spirit as Cherniak, see Hacking (1967), Goldman (1986), Kitcher (1992), Foley (1993), and Pollock (1995, 2006).

\textsuperscript{41}There are many philosophers who, because of the preface paradox or for independent reasons, think that we fallible, human agents are bound to have false beliefs, and perhaps even bound to have inconsistent beliefs. For these people, the preface may not be particularly troubling as an apparent violation of a too-ideal norm of consistency. However, even if you are not troubled by the prospect of widespread inconsistent beliefs, I think the arguments to follow should be of interest. After all, even if it is true that no one has perfectly consistent beliefs, there is still the question of whether sincerely writing a preface like Jennifer’s reveals inconsistent beliefs—that is, there is still a question of whether the preface paradox reveals a genuine inconsistency. I will argue that it does not.

\textsuperscript{42}This sort of “closure” should not be confused with other much-discussed closure principles, e.g., the principle that if one knows that $p$, and knows that $p$ implies $q$, then one is in a position to know that $q$. The latter principle does not entail the one in the text.
the preface case involves Jennifer’s lack of belief in the proposition that all of the claims in the body of her book are true, which follows from her beliefs in each of those claims.

3.2 A Bad Solution to the Preface

One possible line of response to the paradox, of course, is to admit that Jennifer is, indeed, irrational for believing what she says in the body and in the preface of her book, and so are all the many actual writers of pref-aces like Jennifer’s. At first, this might not seem particularly troubling: after all, no one thinks everyone is perfectly rational, and in fact there is a well-established empirical literature showing that people, generally, behave irrationally in various circumstances—cf. the Wason selection task (Wason 1966), Kahneman and Tverksy’s Linda the feminist bank teller (Kahneman et al. 1982), and much of the heuristics and biases literature. And this response has the benefit of allowing us to keep as principles of rationality the appealing consistency and closure constraints. But, as Christensen (2004, §§3.2, 3.3a) makes particularly clear, there is a much greater intuitive cost to counting Jennifer’s preface disclaimer as irrational. For this would require an unjustifiable asymmetry between the way she regards her own fallibility and others’, since there is no reason she would not form a preface-type belief about whether there are errors in another author’s book. Given the same evidence about the likelihood of errors in Brad’s history book—he’s an ordinary, finite agent, writing on a subject that does not admit of deductive proof; his book contains a great many claims; other books like his have always been found to contain at least some errors—Jennifer would conclude that there are at least some errors in the book, and this would not lead to a preface paradox, since Jennifer is not committed to Brad’s claims. But consistency requires a different response to the same evidence about the likelihood of errors in her own history book. There is no non-ad-hoc reason for this asymmetry: Jennifer does not have any special evidence on which it is more likely that Brad’s book contains errors than that her own does. This is particularly clear if we suppose that Brad’s book is on a topic about which
Jennifer knows very little, but on which Brad is an expert (Christensen 2004, §3.3a); in such a case, it makes sense for Jennifer to accept each of Brad’s individual claims as the deliverances of expert testimony. In that case, she has, if anything, better reason to think that she has made a mistake somewhere than that Brad has, since she has a better understanding of the evidence for her own claims and the ways she might have gone wrong somewhere. Of course, the most plausible source of doubts about a book’s infallibility—whether her own or Brad’s—is Jennifer’s awareness that scholarly books of any reasonable length have always been found to have at least some minor errors; and in this regard, her evidence about her own book is exactly the same as her evidence about Brad’s.

Furthermore, if we impose a closure constraint on Jennifer’s beliefs, she must commit to all kinds of intuitively irrational “downstream” beliefs, to use Christensen’s phrase (2004, 49ff). Closure requires that Jennifer not only lack belief in the proposition that there are some errors in her book, but that she in fact believe that her book contains no errors. Coupled with her beliefs about other authors’ books, closure requires that she believe that her book is the first and only error-free history book of its length. If, as we may suppose, she also believes that she has not been significantly more diligent in avoiding possible sources of error in writing her book than other authors in her field generally are, she must also believe that she has been remarkably lucky. She will also believe that her reputation as an historian is about to improve, that she should start seeking better academic positions, that she will soon be driving a better car, and so on. If closure is a constraint on ideal rationality, then Jennifer would be more rational to believe such propositions than not to believe them, which is grossly counterintuitive.

Thus, there is a significant cost to stubbornly holding on to the intuitively appealing principles of consistency and closure.

3.3 A Medium Solution to the Preface

Here is probably the most familiar solution to the preface paradox. So far, the discussion of the preface paradox has only involved outright belief, not
degrees of belief. But once we start thinking in terms of degrees of belief, it becomes obvious why Jennifer writes the preface disclaimer. Since the claims in the body of Jennifer’s book concern contingent facts about the past, if she is rational, her degree of belief in each will be short of certainty, i.e., less than 1. Since she believes each claim to a high enough degree to assert without qualification, though, her degree of belief in each should be fairly high. So the most plausible picture puts Jennifer’s degree of belief in each claim in the body of her book below 1, but close to it. But since a book of any reasonable length is bound to contain a very large number of claims, and since many of those claims can be expected to be probabilistically independent, or nearly so, Jennifer’s degree of belief in their conjunction should be, approximately, the product of her degrees of belief in each claim. The product of a large number of values less than 1 quickly becomes quite small, though, so we should expect that Jennifer’s degree of belief in the conjunction of all the claims in the body of her book should be close to (but greater than) 0. Thus, although she has very high confidence in each of the claims in the body of her book, she has very low confidence in the proposition that they are all true. Equivalently (assuming Jennifer is rational in this regard as well), she has high confidence in the proposition that at least one of her claims is false. Thus, she writes the preface disclaimer.

So far, this is not a solution of the preface paradox (pace Roush 2010, 35–39). I said at the start of this section that until that point, the discussion of the preface had concerned only outright belief, not degrees of belief. This is significant: the statement of the paradox does not have anything to do with degrees of belief; the paradox is a puzzle about outright belief. It is only the more perplexing if appeal to degrees of belief upholds the intuitive verdict that Jennifer is not irrational in writing the book she does, preface and all. If we are to use facts about degrees of belief to solve the preface paradox, then we will need a story about how degrees of belief fit together with outright belief. For example, if outright belief is identical with belief to degree 1, then the above reasoning about degrees of belief does nothing to explain the paradox. For in that case, Jennifer must believe each claim in the body to degree 1, and if her degrees of belief are rational, it follows that her degree of
belief in the preface disclaimer must be 0. So we need an alternative account of the connection between outright belief and degrees of belief if this line of reasoning is to solve the preface paradox.43

The most commonly offered account of the connection between degrees of belief and outright belief is what we may call the threshold view: outright belief is belief to a degree above some threshold value \( t < 1 \). One believes that \( p \), full stop, just in case one’s degree of belief in \( p \) is greater than \( t \). If we add that outright beliefs are rational iff the underlying degrees of belief are rational, then we have enough to complete the solution of the paradox. Jennifer’s high degree of belief in each of the claims in the body of her book, and her high degree of belief in the proposition that at least one of those claims is mistaken, suffice for outright belief in those claims and that proposition; and her degrees of belief are rational, by the reasoning at the start of this section, so the corresponding outright beliefs are rational.

I call this a medium solution to the preface paradox because it is better than the bad solution of §3.2, but worse than my favoured solution, which will appear in §3.4. The main problem with the medium solution is that it relies on the threshold view of outright belief and degrees of belief, which, as I will argue in chapter 4, is deeply flawed. It would be preferable, in fact, to have a solution of the paradox which does not appeal to degrees of belief at all, since degrees of belief do not appear anywhere in the statement of the paradox. We should be able to give an account of what is going on in the case of Jennifer’s book which only appeals to facts about outright belief. This is what I attempt in §3.4.

43One such account one might give is an eliminativist one with respect to outright belief itself, à la Jeffrey (1992). One might claim that this makes for a solution of the paradox: one would say that we cannot account for the preface situation in terms of outright belief, but we can give a satisfactory description purely in terms of degrees of belief, and so much the worse for the notion of outright belief. I will ignore eliminativist strategies here for the reasons I gave in chapter 1: briefly put, this dissertation as a whole can be seen as part of an argument against eliminativism, since I aim to give a coherent and useful account of outright belief.
3.4 A Good Solution to the Preface

3.4.1 Criteria

Before I give my preferred solution, let me say a bit more about what a good solution to the preface paradox would look like. I’ve already said that a good solution must talk about outright beliefs, rather than merely degrees of belief, and that it is better for a solution not to rely on a controversial thesis about the connection between outright belief and degrees of belief.

It is also worth taking a moment to get clear on what a solution must do. At the heart of the paradox is the contradiction between the preface disclaimer and the claims in the body of Jennifer’s book. Accordingly, solving the paradox means resolving a certain apparent contradiction—either by explaining why the contradiction is acceptable, or by explaining why there is no contradiction. My preferred strategy will take the latter route; but if that strategy is to be successful, we must determine what sort of contradiction whose existence one wants to deny.

What I am getting at is that it is much easier to plausibly deny, as I shall, that Jennifer has contradictory beliefs than to deny that she says contradictory things; but it is only philosophically troubling if she has contradictory beliefs. Suppose that Jennifer’s beliefs are entirely consistent, but she has somehow expressed a contradiction in writing her book. It is not clear that this would imply that she is irrational in virtue of having written the book, any more than it is irrational to accidentally write down “\(p \land \neg p\)” instead of “\(p \land \neg q\)”. That would certainly be an error—one has not expressed what one intended to express—but it would not be a manifestation of irrationality. Jennifer, on the other hand, expresses just what she wants to express: she makes certain claims about history, and, in the preface, about her own fallibility.

But it would be worrisome indeed if what she says in the book and its preface indicates that she holds inconsistent beliefs, in part because the mundanity of the preface situation suggests that even those of us who have not written books on history are in the same doxastic predicament. This is the paradox of the preface. To put it another way, the paradox is that Jennifer
has expressed contradictory beliefs, not that she has expressed contradictory propositions. The semantic value of the sentences she writes matters here only insofar as it is reflected in her beliefs. “Most Books Are Sloppy, Have Minor Inconsistencies” is not news, but “Most People Believe Some Minor Inconsistencies” is chilling, at least to people of a certain constitution.

So I take my task to be showing that, given what she says in her book and its preface, Jennifer is not best understood as having expressed contradictory beliefs. I will do this by building on the picture of sincere assertion I put forward in chapter 2. In brief, the solution has it that there is a change of context between the preface and the body of the book, and so the alternatives Jennifer means to rule out in making the preface disclaimer do not conflict with the claims in the body of the book. In broad terms, and ignoring for the moment my sensitivist picture of sincere assertion, this is similar to what Evnine (1999) and Roush (2010) have to say about the preface: both claim that the preface disclaimer is not best represented as the denial of what is claimed in the body of the book.

3.4.2 Intuitive Reasons to Deny Inconsistency

Consider the following scenario. Jennifer has just given the last of a series of lectures, covering the contents of her book. The question and answer period has begun, and Professor Socrates has the first question. Jennifer has been warned that Professor Socrates has an unorthodox way of questioning speakers; since she wants to make a good impression with her hosts, she intends to put up with it.

Socrates: Thank you for a very interesting series of lectures. That was fascinating. But don’t you think you’re being a little bit arrogant?

Jennifer: Excuse me?

S: Well, you’ve made an awful lot of claims. To be fair, I grant that you have excellent evidence for each one, but it’s still unlikely in the extreme that you’d be right about all of those things.
After all, I don’t think anybody has ever produced a work of such length with absolutely no errors along the way.

J: But I never said that all my claims were right, or that I haven’t made any errors at all!

S: You think not? But you just claimed that p, and q, and r, and so on. You’re not retracting any of those, are you?

J: Of course not.

S: So you believe all of those things?

J: Yes, I do.

S: So you think that all of those things are true?

J: I’m pretty sure that’s what I just said. Yes, I think all those things are true.

S: But if all of those things are true—the things you’ve claimed—then that means you haven’t made any errors, right?

J: Well, sure.

S: So you’ve effectively claimed that you’ve made no errors.

Sounds pretty arrogant to me.

So far, I hope Professor Socrates’s objections sound absurd, and unfair.\footnote{In fact, I think Socrates’s objections have the same sort of absurdity as they would if he were presenting standard skeptical objections to a talk about history: “You say that people in fifteenth century Europe had daily lives of a certain sort, but how can you rule out the possibility that there was no fifteenth century, that you are a brain in a vat stimulated to think that there was such a thing as fifteenth century Europe?” It would be unusually patient, but not unreasonable, to respond to such an objection along the lines of Jennifer’s response to Professor Socrates: “My claims should not be taken as saying anything about that possibility. Perhaps, if you press me on it, I might concede that I do not believe my claims with the greatest possible certainty, since I cannot rule out the possibility that I am a brain in a vat. But I need retract nothing to concede this.”}

Before I offer a response on Jennifer’s behalf, note that the reason for this absurdity is not the form of his objections, in the following sense. One might worry that Socrates is assuming some dubious sort of closure principle, according to which, if one has asserted that $p_1, \ldots, p_n$, from which $q$ deductively
follows, then one has asserted that $q$ (perhaps on the additional condition that one recognizes the entailment of $q$ by $p_1, \ldots, p_n$). To be sure, such a principle does indeed sound dubious, and I will not argue that anything of the sort is correct. Nevertheless, people do sometimes, non-absurdly, pose objections in the way Socrates has: “Oh, you say that $p_1, \ldots, p_n$? So you’re saying that $q$!” This is acceptable even when the link between $p_1, \ldots, p_n$ and $q$ falls short of entailment. To take a recent political example, if D has just spoken to R in support of health care reform in the United States, it’s not hard to imagine R responding, “Oh, so you’re saying you believe in socialism?” On the other hand, it is also easy to imagine D responding to R’s claim that government has no business in health care with, “Oh, so you’re saying we should get rid of Medicare and Medicaid too?” Responses like these (a Google search for the phrase “so you’re saying that” turns up many more examples) might be good or bad, but it is not absurd to argue in this way. Socrates’s objections take the same form, but they are absurd.

There are many ways Jennifer might reply.\footnote{In fact, I think Jennifer’s options here parallel Angelina’s options for responding to Brad in the Lunch case of §1.4. That is, by Jennifer’s lights, Socrates has brought up considerations which, technically at least, have some bearing on the things she has been saying, but which she did not mean to say anything about. Just as Angelina did not mean to say anything about whether she would win the lottery when she spoke of her lunch plans for tomorrow, Jennifer did not mean to say anything about whether her lecture series would be error-free when she made her individual claims. However, all that matters for the present is that the line of response in the text, which parallels (LUNCH-ELABORATE), is reasonable.} Perhaps the most plausible reply would be stunned silence, or a polite smile and nod before moving on to the next question. But the following sort of reply is also acceptable.

Jennifer: That’s not right. I’ve made no such claim. I haven’t even addressed the question of whether all my (other) claims are correct; and if I had, I would say that I’m sure I’ve made some errors. But, look, I really haven’t said anything to commit myself to any view about whether I’ve been entirely error-free in my talks. You’re right that there isn’t any one of my claims that I want to retract, and I guess if you really push me on it, I’d have to say I don’t believe any of them without a shadow of a
doubt. Of course I could be wrong about just about any of them, and when you put that many together, I must be wrong about something. But I believe each of these claims about as strongly as I believe anything.

Socrates’s objections are absurd, and Jennifer’s response is reasonable, if unusually patient. The point of this scenario is that there is intuitive pressure to regard the preface disclaimer as consistent with the body claims. (Or, rather, to regard the beliefs revealed by the disclaimer and the body claims as consistent.) This explains the absurdity of Socrates’s objections and the reasonableness of Jennifer’s reply. If the preface disclaimer and the body claims are inconsistent, then Socrates’s argument ought to be hard (at least!) to resist without admitting inconsistency; but Jennifer admits no such thing.

Note that this argument for consistency of the disclaimer with the body claims goes a step further than the arguments in Evnine (1999, 203ff) and Roush (2010, 39–41). Those authors are content to point out that the disclaimer belief is a second-order belief—that is, a belief about one’s beliefs—whereas the claims in the body of the book, in general, are not. Evnine writes:

So we cannot derive [the negation of the conjunction of the body claims] from [the preface disclaimer] unless we make a number of other assumptions. Nor does the recognition of epistemic fallibility directly support [the negation of the conjunction of body claims]. Recognition of one’s fallibility should not be taken to entail an accurate inventory of one’s beliefs. I might well believe that something I believe is false not because I think that one of \( p_1 \ldots p_n \) is false, but because I vaguely think (falsely) that I have some further, unspecified belief that is the culprit.

(1999, 203, emphasis added)

Presumably, among the “number of other assumptions” that must be made to get from the disclaimer to the negation of the conjunction of the body claims would be “an accurate inventory” of the body claims, plus something to the effect that one’s belief that \( p \) is false if and only if \( p \) is false. One
might argue that these assumptions are all that one need add; the following
argument is valid:

1. At least one of the beliefs I espouse in the body of my book is false.

2. $p_1, \ldots, p_n$ is an exhaustive list of the beliefs I espouse in the body of
my book.$^{46}$

3. For any $p$, I have a false belief that $p$ if and only if I believe that $p$ and
$p$ is false.

4. Therefore, at least one of $p_1, \ldots, p_n$ is false.

The extra assumptions 2 and 3 above are, plausibly, ones that an author
might believe to be true. Thus, for at least some plausible authors, the
preface paradox should still arise, despite the fact that the preface disclaimer,
as Evnine and Roush read it, is of a different order than the body claims: if an
author believes 1–3 above, and also believes $p_1, \ldots, p_n$, then she thereby has
inconsistent beliefs. So, in other words, one might worry that Evnine's and
Roush's arguments don't go far enough to establish the lack of contradiction
between the disclaimer and the body claims, since all one need add to the
preface disclaimer in order to derive the negation of the body claims are
some very plausible assumptions.$^{47}$

The argument I offer in this subsection heads off that worry. Professor
Socrates and Jennifer do not disagree on the inventory of claims she has
made, and they do not disagree over what makes a belief false. Nevertheless,
Jennifer insists, reasonably, that she has not committed herself to the denial

---

$^{46}$The version of the preface that Evnine addresses concerns a disclaimer not about the
claims in the body of a book, but rather about all of an agent’s beliefs (other than the
disclaimer belief). It may be less plausible that agents ever have an accurate inventory of
the totality of their beliefs than that authors sometimes have accurate inventories of the
beliefs they espouse in a given book; but Evnine should be worried about this version of
the preface as well as the one he discusses. His aim is to argue that the preface paradox
does not show that accepting the Conjunction Principle (roughly, if one rationally believes
$p$ and rationally believes $q$, then one can rationally believe the conjunction $p \land q$) leads to
allowing rational belief in contradictions. If the version of the preface I am considering
succeeds, involving a book rather than the totality of an agent’s beliefs, then it still follows
that it is sometimes rational to believe a contradiction.

$^{47}$Christensen (2004, 37) makes roughly this point.
of the preface disclaimer (i.e., to the proposition that she has made no errors, no false claims) by committing herself to the claims she has made.

Before we move on, there is an important distinction to make here. The preface paradox is often presented as showing that it is sometimes rational to hold, or that one can sometimes be justified in holding, inconsistent beliefs. (This is how Makinson 1965 originally presented the paradox.) What I think the Jennifer-Socrates dialogue shows, assuming Jennifer’s response to Socrates is reasonable, is that in fact there is no inconsistency between the beliefs indicated by the preface disclaimer and by the claims in the body of her book/lecture series. Her response to Socrates is not that she holds inconsistent beliefs rationally, or justifiedly, but rather that she does not hold inconsistent beliefs—or rather, that she would not be inconsistent in holding the beliefs in question, if she does hold them. (The challenge, of course, is to make sense of this; that is the task of the next subsection.) If, on the contrary, the problem with Socrates’s objections were that he is pointing to an acceptable inconsistency, we should expect a more natural response for Jennifer would be to say something in the spirit of Lewis Carroll’s Tortoise (Carroll 1895):

Jennifer: You’re right. It certainly does follow from what I’ve said that I have made no errors. However, you’re wrong to insist that this makes me arrogant: I don’t believe that I’ve made no errors; on the contrary, I’m quite sure I have made some mistakes. I’m sure that at least one of my claims was false. I acknowledge that this means my beliefs are inconsistent, but I don’t think they’re irrational.

This, I think, is considerably less satisfying as a reply than what I offered previously on Jennifer’s behalf. Regardless of how the two compare, my main point is that the former reply indicates that beliefs corresponding to the preface disclaimer and to the body claims do not contradict one another, and that that reply is reasonable. I bring up the latter reply to make it clear that accepting the beliefs as contradictory but maintaining that one can rationally hold them would be a different thing. Therefore, the reasonableness of the
former reply indicates that there is some intuitive pull to the view that there is no contradiction. In the next subsection, I will argue that my preferred sensitivist model of belief allows us to represent the relevant beliefs as non-contradictory; this constitutes my solution of the paradox.

3.4.3 The Solution, Informally

This subsection gives an informal summary of the one to follow, which is more precise and detailed.

In Chapter 2, we came up with a sensitivist formal framework for modelling belief, and adapted to that model the maxim that an assertion is sincere iff the assertor believes what she asserts. Taking for granted that Jennifer asserts sincerely both the body claims and the preface disclaimer, and also that she has an accurate inventory of the claims in the book, I use the precisified sincerity maxim to draw what conclusions we can about Jennifer's doxastic state. I consider various ways of adapting a consistency norm to the formal framework, and argue that we cannot conclude that Jennifer violates any such norm worth worrying about: in particular, although there is a context where she believes that the body claims are not all true, and there is a context where she believes that they are all true, there need be no context where she believes both that the body claims are all true and that they are not all true. To drive the point home, I give a specific (though simplified) model of Jennifer's doxastic state which would make all of her claims in the body and preface of the book sincere, but which does not manifest any pernicious sort of inconsistency.

3.4.4 The Solution, Formally

In the language of §2.5, we are interested in what Jennifer's assertions in the book tell us about her doxastic state $S$. If we find that there are propositions
p and \(\neg p\)\(^{48}\) such that \(S\) contains\(^{49}\) at least one point where \(p\) and \(\neg p\) are true (i.e., if \(\|p\| \cap \|\neg p\| \neq \emptyset\)),\(^{50}\) then Jennifer’s doxastic state is contradictory, in a certain sense. But this is not quite the right sense for our purposes: that Jennifer’s doxastic state includes some contradictory points, so to speak, does not quite mean that she believes a contradiction, or even that she has contradictory beliefs, unless consistency requires logical omniscience. For recall that we said the population of points in an agent’s doxastic state is determined by the agent’s logical beliefs: for any propositions \(p\) and \(q\), 
\[\|p\| \cap \|q\| = \emptyset\] if and only if the agent thinks that \(p\) and \(q\) are contradictory. So if \(\|p\| \cap \|\neg p\| \neq \emptyset\), this means that Jennifer does not think \(p\) and \(\neg p\) are contradictory, but it does not necessarily mean that there is a context in which Jennifer would believe both that \(p\) and that \(\neg p\). What the preface paradox is meant to show, however, is something stronger than that Jennifer falsely believes some set of jointly inconsistent propositions to be jointly consistent; she is supposed to have inconsistent beliefs about non-logical matters (assuming, let us say, her book was not a logic book).

So we seek a stronger condition on \(S\); but we must be careful not to make the condition too strong, for my aim is to show that the preface paradox does not give an example of an agent with inconsistent beliefs. I must be careful not merely to stipulate Jennifer’s inconsistent beliefs away by making it too difficult to have inconsistent beliefs. Here is a candidate condition, which I think is not too strong. Jennifer’s doxastic state \(S = (U, C, \|\cdot\|, \preceq)\) is inconsistent just in case there is some context \(C \in C\) such that there is a point \(x \in B_C\) with 
\[x \in \|p\| \cap \|\neg p\|\] for some proposition \(p\). That is, Jennifer’s

\(^{48}\)For ease of exposition, I will deal with cases of inconsistent sets of only two propositions. In general, where I write “\(p\) and \(\neg p\),” one can read “\(p_1,\ldots,p_n,\)” and where I write “\(\|p\| \cap \|\neg p\|,\)” one can read “\(\bigcap_{i=1}^n \|p_i\|,\)” all with the understanding that \(p_1,\ldots,p_n\) are jointly inconsistent. No interesting complications arise in the move from two propositions to \(n\) propositions, but the notation gets uglier.

\(^{49}\)This is, of course shorthand for “the universe of \(S\) contains”.

\(^{50}\)Some authors make a distinction between, on the one hand, believing both that \(p\) and that \(\neg p\), and, on the other hand, believing the conjunction of \(p\) and \(\neg p\). For instance, roughly speaking, Kyburg (1970) calls a ban on the former sort of belief set “Strong Consistency,” and a ban on the latter kind of belief “Weak Consistency”. The distinction will make no difference on my account, so I will persist in talking about an agent believing both that \(p\) and that \(\neg p\).
doxastic state is inconsistent just in case there is some context in which she
would fail to rule out an inconsistent set of propositions; in other words, we
might say that there is a context in which she regards an inconsistent set
of propositions as a live possibility. Note that this is weaker than requiring
that she actually outright believe an inconsistent set of propositions. That
would require, not that there be some point \( x \in B_C \) where both \( p \) and \( \neg p \) are
true, but rather that \( p \) and \( \neg p \) be true at all points in \( B_C \). I think it is bad
enough for Jennifer to take a contradictory possibility seriously, failing to rule
it out, even if she is not quite convinced that the contradictory propositions
in question are true.

Now let us see if the preface situation entails that Jennifer’s doxastic
state meets even this condition. To deduce anything about \( S \) from a sincere
assertion of Jennifer’s, we must determine what she took the common ground
to be in the context of the assertion; we must know what possibilities she
meant to rule out. Let \( p_1, \ldots, p_n \) be the claims in the body of the book
(or rather, the propositions expressed thereby). Each \( p_i \) is claimed in some
course \( C_i \). There will surely be some relationship between the various \( C_i \)—
Jennifer’s claims build on each other—but let’s not worry about that right
now. For each \( p_i \), let \( E_i \) be the set of possibilities Jennifer intends to rule
out by claiming that \( p_i \). \( E_i \) will then be a subset of \( C_i \) —in particular, the
subset of \( C_i \) where \( p_i \) is false, \( C_i \cap \Vert \neg p_i \Vert \). Likewise, let \( p_0 \) be the preface
disclaimer, made in the context \( C_0 \), ruling out the possibilities in \( E_0 \). What
we can conclude about Jennifer’s beliefs from the sincerity of her assertions
that \( p_0, \ldots, p_n \), by the formalized sincerity principle of §2.5.2, is that for
each \( C_i \), \( B_{C_i} \cap E_i = \emptyset \). In other words, in the context of each assertion,
Jennifer’s beliefs must rule out all possibilities ruled out by her assertion.
Therefore, for each \( C_i \), there is a set of points \( B_{C_i} \subseteq C_i \cap \Vert p_i \Vert \) such that for
each \( x \in B_{C_i} \) and every \( y \in C_i \), \( x \preceq y \). This is all that we can deduce about
\( S \) from Jennifer’s sincere assertion that \( p_i \) in context \( C_i \).

Let us grant that Jennifer has an accurate inventory of her claims in the
book: she knows that \( p_1, \ldots, p_n \) is an exhaustive list of her claims in the
body of the book. Let us also grant that she regards those propositions,
jointly, as equivalent with their conjunction $\bigwedge_{i=1}^{n} p_i$. She will, thus, be able to infer, from the proposition that one of her claims in the body of the book is false, that one of $p_1, \ldots, p_n$ is false. Let us even suppose that the contexts $C_1, \ldots, C_n$ build on each other monotonically, so that $C_n \subseteq \cdots \subseteq C_1$—each context $C_{i+1}$ is generated from $C_i$ by removing the points where $p_i$ is false, so that $C_{i+1} = C_i \cap \|p_i\|$. In this way, we will have that in $C_n$, $\bigwedge_{i=1}^{n-1} p_i$ is true at all points; so in $B_{C_n}$, $\bigwedge_{i=1}^{n} p_i$ is true. All the suppositions of this paragraph are intended as concessions to the proponent of the preface paradox; I intend these suppositions to make my task harder, not easier.

I have argued that only the preface disclaimer, $p_0$, might plausibly be intended to rule out any possibilities concerning the presence of errors in the book. This is the key to my solution to the preface paradox. At most, the sincerity of the preface disclaimer shows that Jennifer’s doxastic state $S$ (a) contains some points $B_{C_0}$ where at least one of $p_1, \ldots, p_n$ is false, and (b) contains some points $E_0' \subseteq E_0$ where all of $p_1, \ldots, p_n$ are true, such that (c) for every $x \in B_{C_0}$ and every $y \in E_0'$, $x \preceq y$. In other words, given some restriction of the space of possibilities, Jennifer’s doxastic state would rule out all the remaining possibilities where all of $p_1, \ldots, p_n$ are true. We are halfway to meeting the criterion given above for inconsistency of a doxastic state: we have a context where Jennifer believes that $\bigwedge_{i=1}^{n} p_i$, the conjunction of the body claims, is false; now we need only show that in some such context $C$, there is a point in $B_C$ where $\bigwedge_{i=1}^{n} p_i$ is true.

But no such thing will be forthcoming. Whereas the body claims $p_1, \ldots, p_n$ are normally used to establish a belief that would conflict with the impugned preface belief, on the present model what we learn about $S$ from the preface disclaimer does not conflict with what we learn about $S$ from the body claims. For those claims cannot plausibly be regarded as intended to rule out the possibility that there is a false claim in the body of the book. (This was the upshot of Jennifer’s dialogue with Professor Socrates.) True, we have supposed that in $C_n$, Jennifer believes that $\bigwedge_{i=1}^{n} p_i$; but $C_n$ bears no clear
relation to $C_0$. In fact, since none of the sets $E_i$ of possibilities ruled out by the
body claims say anything about whether the book is error-free, we know nothing about the relation between points where $\bigwedge_{i=1}^n p_i$ is true and points where $\bigwedge_{i=1}^n \neg p_i$ is false.

Let’s do some concrete examples. We’ll start simply, with an agent who sincerely asserts two contradictory propositions. This is not quite a preface situation—part of the power of the preface paradox is how common it is for a book to include a preface disclaimer, but it is surely far less common for someone to assert a proposition and its negation (albeit in different contexts). Still, the model to come is instructive, and simpler than one with a larger set of jointly inconsistent propositions.

First, here is a definition. Say a valuation $\parallel \cdot \parallel$ is (propositionally) logically coherent just in case for any propositions $p$ and $q$:

1. $\parallel \neg p \parallel = U \setminus \parallel p \parallel$;
2. $\parallel p \land q \parallel = \parallel p \parallel \cap \parallel q \parallel$; and
3. $\parallel p \lor q \parallel = \parallel p \parallel \cup \parallel q \parallel$.

Say a doxastic state $S$ is logically coherent just in case its valuation $\parallel \cdot \parallel_S$ is logically coherent. If a valuation is logically coherent, then there will be no points at which a contradiction is true, and no points at which a tautology is false.

**Proposition 3.1.** It is possible for an agent to sincerely assert $p$ in one context, sincerely assert $\neg p$ in another context, but not believe both $p$ and $\neg p$ in any context, even if the agent’s doxastic state is logically coherent.

**Proof.** Let $q$ be the proposition that Nixon is a Quaker, let $r$ be the proposition that Nixon is a Republican, and let $p$ be the proposition that Nixon is a pacifist.\(^{51}\) Now let a (confused) agent’s doxastic state be represented by

---

\(^{51}\)The following example is based on the so-called Nixon diamond, which first appears in Reiter and Criscuolo (1981).
$S$, with $U_S$ having eight members, corresponding to the eight possible valuations of the three atomic propositions $q, r, p$. These are pictured in Figure 3.1, labelled according to the valuations of the three propositions. In the diagram, wherever a point $x$ is lower than a point $y$, $x \prec y$; so, for example, we have $qr\bar{p} \prec \bar{q}r\bar{p}$, and $qr\bar{p} \simeq qrp$. Also depicted in the diagram are three members of $C_S$, namely $C_r$, $C_q$, and $C_{total}$. Finally, let the valuations of any non-atomic propositions be such as to make $\|\cdot\|$ logically coherent.

Corresponding to these three context sets, we have three belief sets, $B_{C_r} = \{\bar{q}r\bar{p}\}$, $B_{C_q} = \{q\bar{r}p\}$, and $B_{C_{total}} \{qrp, qr\bar{p}\}$. Thus, in $C_r$, the agent believes $\neg p$ (that Nixon is not a pacifist); in $C_q$, the agent believes $p$ (that Nixon is a pacifist); and in $C_{total}$, the agent believes neither $p$ nor $\neg p$. Therefore, in a context corresponding to $C_r$, the agent could sincerely assert that $\neg p$, and in a context corresponding to $C_q$, the agent could sincerely assert that $p$.

This is meant to be a fairly natural case. We have an agent who thinks that Quakers are usually pacifists, Republicans are usually not pacifists, and is unsure, all things considered, whether Nixon is a pacifist or not. $C_r$ is meant to correspond to a context in which it is particularly vivid for the agent that Nixon is a Republican, and accordingly not a pacifist; $C_q$ is meant
to correspond to a context in which it is particularly vivid that Nixon is a Quaker, and accordingly a pacifist; and $C_{total}$ to a more synoptic context, where the agent realizes that Nixon is both a Quaker and a Republican, and the agent is accordingly unsure whether Nixon is a pacifist. If the agent has not thought much about $p$, $q$, or $r$ before, arguing persuasively that Nixon is a Quaker and therefore must be a pacifist, perhaps with lots of specific evidence of Nixon’s Quakerhood and of Quakers’ pacifism, might suffice to put the agent into the context $C_q$; likewise for $C_r$.

Now let us look at a more complicated model that represents the preface situation. Let $S_{pref} = (U_{pref}, C_{pref}, \parallel \cdot \parallel_{pref}, \preceq_{pref})$ be a doxastic state for the propositions $p_1, \ldots, p_n, q_0, q_1, \ldots, q_n$. (I will suppress the subscripted “pref” when doing so will not cause confusion.) Let the members of $U$ be $a_1, \ldots, a_n, b_0, b_1, \ldots, b_n, c_1, \ldots, c_{2^n-1}$. For $1 \leq i \leq n-1$, let $a_{i+1} \prec a_i$ and $b_{i+1} \simeq a_i$; let $b_0 \prec a_n$; and for $1 \leq i, j \leq 2^n-1$, let $c_i \prec b_0$ and $c_i \simeq c_j$.

A basic diagram of $U$ and $\preceq$ is given in Figure 3.2a. In the diagram, if $x$ appears lower than $y$, then $x \prec y$; I have added arrows from $y$ to $x$ to highlight the cases where it is important that $x \prec y$.

For $2 \leq i \leq n$, let $E_i = \{a_{i-1}, b_i\}$, and let $E_1 = \{b_1\}$ and $E_0 = \{b_0\}$ (Figure 3.2b). Let $C$ include the sets $C_0, C_1, \ldots, C_n$, defined as follows. Let $C_1 = \{a_1, \ldots, a_n, b_1, \ldots, b_n\}$; for $2 \leq i \leq n$, let $C_i = C_{i-1} \setminus E_{i-1}$; and let $C_0 = \{b_0, c_1, \ldots, c_{2^n-1}\}$. These are depicted in Figure 3.2c. Note that $\{C_0, C_1\}$ partitions $U$.

Finally, we must give valuations to the propositions in question. Let $\parallel \cdot \parallel$ be logically coherent. For $0 \leq i \leq n$, let $\parallel q_i \parallel = E_i$. For $1 \leq i \leq n-1$, let $\parallel p_i \parallel \cap C_1 = C_{i+1}$; and let $\parallel p_n \parallel \cap C_1 = \{a_n\}$. Now all that remains is to give the values of $p_1, \ldots, p_n$ on $C_0$. First, the values of $p_1, \ldots, p_n$ on $b_0$ are unimportant for our purposes; the case where they are all true at $b_0$ has a special interpretation which I will return to in the discussion of the model. Otherwise, choose arbitrary values for $p_1, \ldots, p_n$ at $b_0$. Now we only have $c_1, \ldots, c_{2^n-1}$ left to deal with. Note that there are $2^n$ possible distributions of truth values among the $p_1, \ldots, p_n$. Give those distributions some ordering from 1 to $2^n$, where the $2^n$th distribution is the one where all of $p_1, \ldots, p_n$ are true. Then for $1 \leq i \leq 2^n-1$, let the values of $p_1, \ldots, p_n$ at $c_i$ correspond
Figure 3.2: $S_{pref}$: (a) The universe $U$ and ordering $\preceq$; (b) the contrasts $E_i$; (c) the contexts $C_i$. 
to the $i$th distribution. Thus, $c_1, \ldots, c_{2^n-1}$ cover all the possible ways there might be a false proposition among $p_1, \ldots, p_n$.

Now let us check what the agent believes in each context $C_0, \ldots, C_n$. For $1 \leq i \leq n$, $B_{C_i} = \{a_n\}$, and so $B_{C_i} \models \bigwedge_{j=1}^{n} p_j$, and $B_{C_i} \models \bigwedge_{j=0}^{n} \neg q_j$. On the other hand, $B_{C_0} = \{c_1, \ldots, c_{2^n-1}\}$; and for every $1 \leq i \leq n$, there is a $1 \leq j \leq n$ such that $c_j \models p_i$, and a $1 \leq k \leq n$ such that $c_k \models \neg p_i$. Therefore, in $C_0$, the agent does not believe that $p_i$ or $\neg p_i$ for any $1 \leq i \leq n$. However, $B_{C_0} \models \neg \bigwedge_{i=1}^{n} p_i$, and $B_{C_o} \models \bigwedge_{i=0}^{n} \neg q_i$.

That is the model; let me explain why it allows the preface situation. The claims in the body of the book, $p_1, \ldots, p_n$, are respectively asserted in the contexts $C_1, \ldots, C_n$. Each assertion $p_i$ is intended to rule out the possibilities in $E_i$, resulting in the updated context set $C_{i+1}$.

These assertions are sincere, since for each $i$, $B_{C_i} \models p_i$. The preface disclaimer (for the sake of argument, let the disclaimer be $\neg \bigwedge_{i=1}^{n} p_i$, the proposition that at least one of $p_1, \ldots, p_n$ is false) is made in the context $C_0$. The author intends to rule out the possibility that $q_0$ is true, leaving only points where at least one of $p_1, \ldots, p_n$ is false. Again, the preface disclaimer is sincere, since $B_{C_0} \models \neg \bigwedge_{i=1}^{n} p_i$.

What do the various propositions $q_i$ mean? Well, $q_1, \ldots, q_n$ will have something to do with the author’s evidence for the individual body claims $p_1, \ldots, p_n$. The case of $q_0$ is a bit more complicated. If $b_0 \models \neg \bigwedge_{i=1}^{n} p_i$, then in the context $C_0$ it is taken for granted that at least one of the body claims is false: $C_0 \models \neg \bigwedge_{i=1}^{n} p_i$. So $q_0$ must be compatible with this presumption. Perhaps it will say that at least one of the people Jennifer thanked is responsible for some of the errors; then ruling out points where $q_0$ is true would leave uneliminated only points where Jennifer is solely responsible for the book’s errors. It is plausible that this is what is meant by a disclaimer of the familiar form, “All remaining errors are my fault, not theirs.” On the other hand, if $b_0 \models \bigwedge_{i=1}^{n} p_i$, then $q_0$ could (but need not) be the proposition that there are

---

52 Except, of course, for the final claim, that $p_n$, since we did not define $C_{n+1}$. But the main point stands: $p_n$ is asserted in $C_n$ with the intention of ruling out $E_n$. 90
no errors in the book.

Thus, \( S_{\text{pref}} \) represents the preface situation without committing the believer to any undesirable sort of inconsistency, or attributing to her any implausible beliefs. The paradox is solved.

### 3.5 Epilogue: Rationality and Consistency

This chapter might, at first glance, seem out of place in this dissertation. After all, I have been careful to insist that my overall project here concerns belief, not rational belief; but the preface paradox is a paradox about rationality. The paradox is that it is intuitive that inconsistent belief is irrational, but it is also intuitive that the preface writer expresses inconsistent beliefs and that such prefaces are innocuous. My solution to the paradox ignores the debate over whether consistency is a rational requirement, and instead argues that the preface writer does not express inconsistent beliefs after all. (She may have inconsistent beliefs, but her sincerely writing the preface disclaimer does not show it.) Therefore, we can conclude that the preface situation is irrelevant to the debate over whether consistency is a rational requirement.

But there is a somewhat broader upshot for that debate, as well. We saw in §3.4.4 that it was not perfectly clear how to define inconsistency for a doxastic state, provided we want inconsistency to be significantly worse than merely having mistaken logical beliefs. If I am right about how epistemologists should think of beliefs—that is, as things with the same structure as the doxastic state models introduced in §2.5.2—then the debate over whether consistency is a rational requirement needs careful framing. My own opinion, for which I will not argue here, is that there is no interesting question left about whether consistency is a rational requirement.

This chapter concludes the portion of the dissertation in which we consider only outright belief. In the next chapter, I will offer an extension of the present model to deal with degrees of belief as well as outright belief.
Chapter 4

Degrees of Belief

4.1 Introduction

In this chapter, I argue for the following two theses:

(CONTEXT) Degrees of belief change from context to context, depending on the space of alternative possibilities.

(UNITY) Outright belief is belief to degree 1.

These theses constitute an extension of sensitivism to partial belief, in the sense that, taken together, they imply sensitivism about outright belief. There are, of course, other ways one might extend sensitivism to partial belief—(CONTEXT) by itself, without any claim about the relation between partial belief and outright belief, could even be considered such an extension, despite not entailing sensitivism about outright belief. For simplicity, though, I will henceforth refer to the conjunction of (CONTEXT) and (UNITY) as sensitivism about degrees of belief, or credence-sensitivism. (Degrees of belief are sometimes referred to as credences.) By contrast, I will refer to sensitivism about outright belief—the topic of the preceding chapters—as belief-sensitivism.

The case for credence-sensitivism does not quite parallel the case for belief-sensitivism. In the case of belief-sensitivism, I pointed to suggestive empirical work and described a dialogue that would be hard to make sense of without sensitivism (chapter 1), then argued that we need sensitivism to make sense of sincere assertion (chapter 2) and applied a sensitivist model of belief to resolve the preface paradox (chapter 3). This chapter will have much more in common with chapter 3 than with chapters 1 and 2: I will describe a sensitivist way of modelling degrees of belief, and apply that model
to solve problems. The usefulness of this model in solving those problems constitutes my case for credence-sensitivism.

In particular, I will claim (§4.3.1) that (UNITY) provides a nice resolution of Kyburg’s (1961) lottery paradox, and (§4.4) that (CONTEXT) resolves the usual problems with (UNITY). In §4.5, I describe my proposed way of modelling degrees of belief, and show in §4.5.1 that it generalizes the way of modelling outright belief set out in chapter 2. The way of modelling degrees of belief I will describe in this chapter can be seen as a modification of the subjective Bayesian approach to degrees of belief, so, finally, I will argue in §4.6 that accepting credence-sensitivism does not threaten the usefulness of modelling degrees of belief probabilistically: the major triumphs of Bayesianism still work on my sensitivist revision of the Bayesian framework, and my version of Bayesianism has some advantages over orthodoxy.

4.2 Orthodox Formalism

I have already let on that the way of modelling degrees of belief I favour is a modification of the standard subjective Bayesian framework. It makes sense, then, to begin with a brief summary of the Bayesian framework.

An agent's degrees of belief are modelled by a personal probability function, \( \Pr(\cdot) \). The function takes propositions as arguments, and returns real numbers between 0 and 1 as values. \( \Pr(p) = 1 \) means the agent has the highest confidence in \( p \), or is certain that \( p \) is true; \( \Pr(p) = 0 \) means the agent is certain \( p \) is false; and \( \Pr(p) = 0.5 \) means the agent thinks \( p \) is as likely to be true as to be false. Values in between 0.5 and 1 indicate that the agent takes \( p \) to be more likely true than false, with increasing degrees of strength, and the case is symmetrical for values between 0.5 and 0. Because \( \Pr(\cdot) \) is a probability function (i.e., it obeys the axioms of the probability calculus), we know a few things about the degrees of belief of an ideal Bayesian agent: for instance, every tautology gets credence 1 and every contradiction gets 0; any two logically equivalent propositions get the same credence; if \( p \) entails \( q \), then \( q \) gets at least as high a credence as \( p \); and so on.

So much for statics. Diachronically, a Bayesian agent revises her degrees
of belief only when she gains new evidence, and does so through the process of *updating by conditionalization*. If she learns some new evidence proposition, \( e \), then she replaces her old personal probability function \( \Pr_{\text{old}}(\cdot) \) with a new one, \( \Pr_{\text{new}}(\cdot) \), defined by

\[
\Pr_{\text{new}}(p) = \Pr_{\text{old}}(p|e),
\]

where \( \Pr(a|b) \) is the probability of \( a \) conditional on \( b \), defined as \( \frac{\Pr(ab)}{\Pr(b)} \), provided \( \Pr(b) \neq 0 \). Two consequences of this updating rule are worth pointing out. First, if \( \Pr_{\text{old}}(e) = 0 \), then \( \Pr_{\text{new}}(p) \) is undefined. So once one has assigned probability 0 to a proposition, one cannot learn that it is, after all, true. Second, if \( \Pr_{\text{old}}(p) = 1 \), then \( \Pr_{\text{new}}(p) = 1 \) as well: \( \Pr_{\text{old}}(p|e) = 1 \), for any proposition \( e \) with \( \Pr_{\text{old}}(e) \neq 0 \). So once one has assigned a proposition probability 1, one’s credence in it can never drop any lower. This should seem, at first, to be a problem for (UNITY); in §4.4, I shall use (CONTEXT) to offer a way out.

This is what degrees of belief are like. How do we fit outright belief into the picture? One prominent view, which I’ll call the *threshold view*, has it that outright belief is belief to a degree higher than some threshold value \( x < 1 \): one believes that \( p \) (outright) if and only if \( Cr(p) \geq x \), where \( Cr(\cdot) \) is one’s credence function. (UNITY) can, of course, be seen as a limiting case of the threshold view: if the threshold \( x \) is set to 1, then the threshold view coincides with (UNITY). However, there are qualitative differences between the views that result from setting the threshold at 1 and setting the threshold below 1, so it is best not to think of (UNITY) as a threshold view.

The threshold view is often coupled with a view about rational belief which I’ll call the *Lockean view*, following Foley (1993), according to which

---

53There are, of course, ways of dealing with this problem, e.g., by taking conditional probabilities as primitive and defining unconditional probabilities in terms of them, rather than the other way around (Popper 1959). See Hájek (2003) for a compelling argument against taking \( \frac{\Pr(ab)}{\Pr(b)} \) as a *definition* of \( \Pr(a|b) \) rather than a *constraint* on \( \Pr(a|b) \) in cases where \( \Pr(b) \) is defined and non-zero. However, there is no consensus on how best to deal with this problem; even if one takes conditional probabilities as primitive, there is no consensus on which axiomatization of conditional probability is correct. See also note 65 on page 111.
one is rational to believe $p$ outright if and only if one is rational to have degree of belief that $p$ higher than some threshold value $y$.\footnote{This can be read in either a permissive or an obligatory way: one is rationally required/permitted to believe that $p$ iff one is rationally required/permitted to have $Cr(p) > y$. Bayesian orthodoxy has it that rational degrees of belief are precise, and so (once one’s prior credence function is fixed) exactly one credence is rationally permitted, and hence also required. (Subjective and objective Bayesians differ over whether there are multiple permissible prior credence functions.) There are a number of authors who have argued that credences are best represented by sets of probability functions, but most of them do not discuss the relation between outright belief and degrees of belief (or at least, not in the same place). See, e.g., van Fraassen (2006), Joyce (2005), Levi (1980), and Weirich (2001). Sturgeon (2008), notably, endorses the threshold and Lockean views, but thinks that degrees of belief are best represented by sets of probability functions; so for him, there is a real difference between what credences are rationally permissible and what credences are rationally obligatory.} To be perfectly clear: the threshold view has to do with the relationship between outright belief and degrees of belief; the Lockean view has to do with the relationship between rational belief and rational degrees of belief. Generally, authors who endorse both the Lockean view and the threshold view, as Foley (1993) does, take it that the two thresholds, $x$ and $y$, coincide: that is, that one is rational to believe that $p$ outright just in case one is rational to have a degree of belief high enough to constitute or entail outright belief. However, one could take both the threshold view and the Lockean view without identifying the two thresholds. On such a combination of views, it would be possible for rationality to demand one have $x < Cr(p) < y$ (if $x < y$). One would then have precisely the degree of belief in $p$ required by rationality, and this would entail that one believe outright that $p$ because $x < Cr(p)$; but one’s outright belief would be irrational, because $Cr(p) < y$. Or, if $y < x$, it would be possible that one have precisely the degree of belief in $p$ required by rationality, and this would entail that one not believe outright that $p$ because $Cr(p) < x$; but one would be rational to believe that $p$. These are peculiar enough possibilities to make the identification $x = y$ quite appealing.

There are some authors who endorse the Lockean view without endorsing the threshold view. This, I think, is the best way to understand Hawthorne and Bovens (1999), which aims to derive rules for rational belief from the Lockean view plus probabilism on degrees of belief. That is, Hawthorne and
Bovens aim to answer the question, “Given my degrees of belief, what ought I to believe outright?” But this question makes no sense if one’s degrees of belief determine one’s outright beliefs, as the threshold view would have it. On the other hand, I know of no one who endorses the threshold view without endorsing the Lockean view; but, of course, this should not be taken to mean that the one view entails the other.

4.3 Problems for the Threshold View

4.3.1 The Lottery

Here is a problem for the threshold view, based on Kyburg’s (1961) lottery paradox.

You own a lottery ticket. You know that there are $n$ tickets in the lottery, that exactly one winner will be selected, that each ticket has an equal chance of winning, and that each ticket’s winning or losing is independent of each other ticket’s winning or losing. You have credences corresponding to these propositions: for any ticket $i$, let $W_i$ be the proposition that $i$ wins, and $L_i$ the proposition that $i$ loses; then for all $i$, you have $Cr(L_i) = 1 - Cr(W_i) = \frac{n-1}{n}$. Furthermore, for any set of tickets $S$ with $m$ members, you have $Cr(L_S) = \left(\frac{n-1}{n}\right)^m$, where $L_S$ is the proposition that all of the tickets in $S$ lose. These credences simply reflect your (accurate) beliefs about how the lottery is set up.

Now here is the problem: is it possible for you not to believe that your ticket will lose? On the threshold view, if $n$ is sufficiently high, this is impossible. For whatever the value of the threshold of belief, $x$, there will be some $n$ such that $\frac{n-1}{n} \geq x$. This is strikingly unintuitive. Certainly there are cases of people who fail to believe that certain particular tickets (their own or others’) will lose what they believe to be large, fair lotteries; it is hard to make sense of the idea that none of these people can have credences that accurately reflect their beliefs about the fairness and size of the lotteries in question.

---

55 Weatherson (2005) makes essentially the same point.
Note that the version of the paradox here is distinct from a related version of the paradox widely discussed in the literature (and which is closer to Kyburg’s original formulation of the paradox). I have in mind the version of the paradox used to make trouble for the Lockean view, possibly but not necessarily in combination with the threshold view. For clarity, let us refer to the two versions of the paradox as the threshold lottery paradox and the Lockean lottery paradox, respectively. Whereas the threshold lottery paradox described above is a problem for certain accounts of belief, the Lockean lottery paradox is a problem for certain accounts of rational belief. Here, the problem would be as follows. For each ticket $i$, your knowledge that the lottery is fair makes it rational for you to have $Cr(L_i) = \frac{n-1}{n} \geq y$, where $y$ is the threshold for rational belief. Thus, you are rational to believe $L_i$ outright, for each ticket $i$. However, you are also rational to believe that exactly one ticket will win; thus, you are rational to believe that at least one of the $L_i$ is false. Given some additional assumptions\footnote{Notably, the assumption that it is rational to believe the conjunction $p \land q$ if it is rational to believe $p$ and rational to believe $q$ often features in the presentation of the problem. However, if this assumption is correct, and if rational degrees of belief satisfy the probability calculus, then the Lockean view must be false. For any two independent propositions $p$ and $q$, the rational agent will have $Cr(p \land q) = Cr(p) \cdot Cr(q)$, which will be lower than either $Cr(p)$ or $Cr(q)$ if neither is believed to degree 1; thus, if $Cr(p)$ and $Cr(q)$ are only slightly above the threshold $y$, then $Cr(p \land q)$ will be below $y$. Thus, we would have a case where it is rational to believe $p$ and to believe $q$, but not rational to believe their conjunction $p \land q$. However, rejecting the above assumption is itself an intuitive cost to the Lockean view. Furthermore, it is not necessary to rely on this assumption to run the paradox. For example, one can massage the principle of noncontradiction so that it rules out rationally believing all the members of an inconsistent set, instead of just ruling out rational belief in a logically false proposition.} it follows that you are rational to believe a contradiction. This conflicts with the intuitive principle of noncontradiction. One response to this problem, taken by Foley (1993) and Hawthorne and Bovens (1999), is to reject the principle of noncontradiction.

The threshold lottery paradox is more difficult than the Lockean lottery paradox. It will not do simply to reject the principle of noncontradiction, or any other principle of rationality for that matter. The problem has nothing to do with what beliefs are rational for an agent; rather, it has to do with
what beliefs are even possible for an agent. According to the threshold view, if you believe that the lottery is fair and that it is sufficiently large, and if your degrees of belief reflect this, then you must believe of each ticket that it will lose. According to the threshold view, it is not merely irrational, but impossible, not to believe that your lottery ticket will lose. An advantage of the threshold view over eliminativism (cf. §1.3) is that it holds on to belief-talk—it counts talk about outright belief as legitimate. This is an advantage because much of our ordinary and theoretical reasoning about knowledge and rationality appeals to outright belief. But the threshold lottery paradox shows that this advantage of the threshold view is overblown, for the threshold view forces a radically revisionist view of some of our belief-talk: it turns out that nobody ever believes that she owns a ticket which might win a fair lottery of any size. This is still not quite so bad as saying that all belief-talk is illegitimate, or must be reinterpreted as credence-talk, but the difference is one of degree, not of kind.

4.3.2 Other Problems

In this subsection, I will briefly present some other (at least prima facie) problems for the threshold view which have arisen in the literature, and argue they are not problems for (UNITY). I will not attempt to push any of these problems as a decisive objection to the threshold view—in fact, in one case I argue that the supposed problem is no problem at all for the threshold view—but I think it is an advantage to be able to avoid easily a number of problems which have been put forward in print.

First, note that there is a clear qualitative difference between belief and lack of belief, but there is no clear qualitative difference between, say, credence 0.9001 and credence 0.8999. If the threshold for belief is 0.9, then the former credence amounts to outright belief, but the latter does not. To put the problem another way, here is Stalnaker (1984, 91):

One could easily enough define a concept of acceptance.\textsuperscript{57}

\footnote{57For Stalnaker, belief is a species of acceptance, though there are other species, such as assumption, supposition, and presumption. The difference between belief and acceptance}
which identified it with a high subjective or epistemic probability (probability greater than some specified number between one-half and one), but it is not clear what the point of doing so would be. Once a subjective or epistemic probability value is assigned to a proposition, there is nothing more to be said about its epistemic status. Bayesian decision theory gives a complete account of how probability values, including high ones, ought to guide behavior, in both the context of inquiry and the application of belief outside of this context. So what could be the point of selecting an interval near the top of the probability scale and conferring on the propositions whose probability falls in that interval the honorific title “accepted”? Unless acceptance has some consequences, unless the way one classifies the propositions as accepted, rejected, or judgment suspended makes a difference to how the agent behaves, or ought to behave, it is difficult to see how the concept of acceptance can have the interest and importance for inquiry that it seems to have.

If we know an agent’s credences, it is not clear what point there is to asking the further question of what the agent believes outright. Any question that might be answered using information about the agent’s outright beliefs could be answered as well or better using information about the agent’s credences, if the threshold view is correct. If we take it for granted, as I think we should, that belief-talk is genuinely useful and important, then this is a problem for the threshold view: the threshold view threatens to render belief-talk pointless.

I think it is clear that (UNITY) does not have this problem. There is a qualitative difference between believing $p$ to degree $1 - \epsilon$ and believing it to degree 1: it is the difference between having some doubt that $p$ and having no doubt that $p$. One who believes $p$ to degree 1 will not regard a bet on $p$ as risky; this is not so for one who believes $p$ to degree $1 - \epsilon$. To be sure, if $\epsilon$ is small enough, the agent may wind up making all the same $p$-related

\hspace{1cm}

is irrelevant to the point being made in the quoted passage.
decisions, but one reasons differently under risk than otherwise. One who believes \( p \) to degree 1 will not seek further evidence for or against \( p \), but one who believes \( p \) to degree \( 1 - \epsilon \) will do so, provided the cost of seeking evidence is not too great and the prospects of finding such evidence are good enough. There is a clear point to distinguishing the propositions an agent believes to degree 1; belief-talk is not pointless if (UNITY) is true.

There is another qualitative difference between credence 1 and any lesser credence, of particular interest to Bayesians: if you give \( p \) credence 1, then you will never give \( p \) any lesser credence if your credences change only via conditionalization, as we saw in §4.2. The same is not true for any credence lower than 1, except for 0. However, I cannot wholeheartedly endorse this argument, because the status of updating via conditionalization is less clear on my view than it is on the orthodox Bayesian view; see §4.6 for more on this point.

Second, Fantl and McGrath (2009, 141) list the following “apparent facts about outright belief”:

- **The truth standard.** If you believe \( p \) and \( p \) is false, then you are mistaken about whether \( p \), and if you believe \( p \) and \( p \) is true, then you are right about whether \( p \).

- **Transparency.** Strength of justification for \( p \) is directly proportional to the strength of justification you have for believing \( p \).

- **Resolution.** If you believe that \( p \), then your mind is made up that \( p \).

Fantl and McGrath then go on to argue that the threshold view cannot account for these apparent facts; I will consider each of their arguments, and conclude that given (UNITY), we can account for all three apparent facts.

- **The truth standard.** The trouble for the threshold view here is that giving \( p \) high credence does not seem to mean commitment to the truth of \( p \). If one’s credence in \( p \) is 0.99, then one thinks that \( p \) is very likely, but one allows that \( p \) might be false. As Fantl and McGrath write: “If you are told: ‘Ha, so you were wrong about whether \( p \), weren’t you?’ you could reasonably...”
say in your defense: ‘Look, I took no stand about whether $p$ was true or not; I just assigned it a high probability; I assigned its negation a probability, too’” (2009, 141).

On the other hand, if one believes $p$ to degree 1, and $p$ turns out to be false, it is hard to see how one could reasonably claim not to have been wrong about $p$. There was no alternative to $p$ to which one assigned a non-zero probability; one did not take not-$p$ to be a live possibility. In particular, the response Fantl and McGrath offer above is not available.

Fantl and McGrath also point out that in high-stakes situations, agents who give $p$ high credence may nevertheless say things like “I don’t want to be wrong that $p$,” and refuse to rely on $p$ in their actions. They claim that this is “some evidence that ordinary folk are thinking of a belief state that isn’t [threshold] belief but is more closely tied to action” (2009, 142). I take it the idea is that these agents refuse to rely on $p$ in their actions because to do so when $p$ is false would mean that they are wrong that $p$; and this is the same sort of being wrong about $p$ which applies to outright belief that $p$.

I find this argument unconvincing. First, it seems to me there is a clear sense of being wrong that $p$ or about $p$ which has nothing to do with belief. One who performs an action which is beneficial or optimal if and only if $p$ is true was wrong to do so if $p$ is false, in a certain sense; and this is still the case even if the agent did not believe that $p$. For example, suppose the A train would take me where I want to go, and the B train would take me away from where I want to go. Suppose I have 60% confidence that the train in front of me is an A train, and 40% confidence that it is a B train. I certainly do not believe that the train is an A train; I have not formed an outright belief. Nevertheless, if I get on the train, and it turns out to be a B train, it is natural to describe me as having been wrong about whether it was an A train, or wrong that it was an A train. It does not matter to this sense of being wrong what my beliefs are.

58 Though, to be fair, Fantl and McGrath only claim that there is “some evidence” for an action-centric notion of belief here. It is possible that the following only answers a bolder argument than the one they mean to endorse; nevertheless, I think it is worthwhile to explore alternative ways of explaining the phenomenon in question.
Transparency. The crucial claim in Fantl and McGrath's argument here is that something might justify one in increasing one's credence that \( p \) without giving any justification for having a high credence that \( p \). But according to the threshold view, believing that \( p \) outright is the same as having a high credence that \( p \); so something might justify increasing one's credence that \( p \) without giving any justification for believing that \( p \). This violates Transparency, provided that giving justification for increasing one's credence that \( p \) entails giving justification for \( p \). The example Fantl and McGrath give is of buying a lottery ticket: buying a lottery ticket in a large lottery, they say, slightly increases your justification for believing that you will win the lottery; however, they claim that it does not give any justification at all for having a high credence that you will win the lottery. Note that this argument weighs against (UNITY) as well as the threshold view, for credence 1 is a species of high credence.

I think the argument from Transparency is too quick. Plausibly, if one's evidence requires one to have credence \( x \) that \( p \), then one is more justified by one's evidence in having credence \( x + \epsilon \) in \( p \) than in having credence \( x + \delta \) in \( p \) if \(|\epsilon| < |\delta|\). That is, if one's credences deviate further from the rationally required or optimal credence, then one's credences are less justified. It follows, then, that if one receives some evidence that justifies raising one's credence in \( p \) by some amount, then that evidence also justifies one in having a high credence that \( p \), although this may not be a very high degree of justification. If one buys a ticket in a fair lottery of 1,000 tickets, one thereby comes to be (fully) justified in having a credence of 0.001 that one will win the lottery (provided one keeps one's ticket, etc.). But one also gains some justification for having a credence of 0.9 that one will win the lottery: whereas before buying the ticket, such a credence would differ from rationality by 0.9, now it differs from rationality by only 0.899. This is a small increase in justification, to be sure, but it is an increase.

Of course, the argument of the preceding paragraph relies on a certain view of justification of credences for which I have not argued. But my point here is that Fantl and McGrath need to supply some conflicting view of justification of credences in order for their objection to the threshold view to
go through. It is not obvious to me that this can be done unproblematically.

*Resolution.* Here it is fairly clear what the problem is for the threshold view, and why there is no problem for (UNITY). An agent who has a high, but sub-unity, credence that \( p \) may well do things like seek further evidence that \( p \), and refuse to act on the assumption that \( p \) without further evidence. (This is particularly likely in high-stakes cases.) It is hard to describe such an agent as having her mind made up that \( p \). On the other hand, if an agent gives \( p \) credence 1, she will not engage in any such activities, since there are no live not-\( p \) possibilities for her.

### 4.4 (CONTEXT) Makes (UNITY) Plausible

It is standardly held that (UNITY) puts too strict a condition on outright belief: if outright belief requires the highest possible degree of belief, it is thought, then ordinary people rarely or perhaps never have any outright beliefs, and rational people might only believe tautologies outright. Here are some representative passages:

Indeed, it might not even matter much where the threshold is as long as we are consistent in applying it. There are some restrictions, of course. We won’t want to require subjective certainty for belief. So, the threshold shouldn’t be that high. . . . Suppose that degrees of belief can be measured on a scale from 0 to 1, with 1 representing subjective certainty. Let the threshold \( x \) required for belief be any real number less than 1. . . . [W]e have already agreed that \( x \) need not be 1. Subjective certainty is not required for belief. (Foley 2009, 38–9; cf. Foley 1993, 143–4)

What is the relation between acceptance and probability? One suggestion would be to identify acceptance of a hypothesis with assignment of probability 1 to that hypothesis. But this view is untenable. For to give hypothesis \( H \) probability 1 is to be willing to bet on it at any odds; for example, a person who gave \( H \) probability 1 would be willing to accept a bet in which the person
wins a penny if $H$ is true, and dies a horrible death if $H$ is false.

I think it is clear that scientists are not usually this confident of the hypotheses they sincerely categorically assert, and thus that probability 1 is not a necessary condition for acceptance.

(Maher 1993, 133)

The usual view seems to be that you do not need to be absolutely certain of $H$ (give it probability 1) in order to believe it. For one thing, it is usually supposed that there is very little we can be rationally certain of, but that we can nevertheless rationally hold beliefs on a wide range of topics.

(Maher 1993, 152)

The Certainty View [is as follows:] You count as believing $p$ just if you assign $[Cr(p)]$ the value 1. . . . [T]he Certainty View takes belief to entail so great a commitment that (i) for each of your beliefs, you must be prepared, for no gain, to gamble any sum that it is true, (ii) you must be prepared to make a like gamble, for each hypothesis that is incompatible with one of your beliefs, that this hypothesis is false, and (iii) you must be prepared to make a like gamble that you haven’t even one false belief. Given this much, it is hard to see how there can be very much you have any business believing. The conclusion is inescapable. If, in focusing on what we are justified in believing (rational to believe), traditional epistemology is concerned with anything at all central to our lives as inquirers, then the Certainty View must be mistaken. (Kaplan 1996, 91, 93)

[(UNITY)] is, I think, less plausible [than the threshold view]. If the binary conception of belief derives its plausibility from our habit of making unqualified assertions, and from our ordinary ways of thinking and talking about belief, then the plausible notion of binary belief is of an attitude that falls far short of absolute certainty. We often assert, or say that we believe, all kinds of things of which we are not absolutely certain. This is particularly
clear if the plausibility of the graded conception of belief is rooted in part in how belief informs practical decision. Insofar as degree of belief is correlated with practical decision-making, the highest degree of belief in P is correlated with making decisions that completely dismiss even the tiniest chance of P’s falsity. For example, having degree of belief 1 in Jocko’s having cheated would correlate with being willing literally to bet one’s life on Jocko’s having cheated, even for a trivial reward. Surely this level of certainty is not expressed by ordinary unqualified assertions; nor is it what we usually want to indicate about ourselves when we say, e.g., “I believe that Jocko cheated,” or what we want to indicate about others when we say, e.g., “Yolanda believes that Jocko cheated.”

(Christensen 2004, 21)

Consider first the suggestion that flat-out belief is maximum confidence—a view reflected in the frequent use of the term “full belief” for flat-out belief. The problem here is that one can believe something, in the everyday sense, without being certain of it. I believe that my grandmother was born on the 3rd of August, but I am not absolutely certain of it. I may have misremembered or been misinformed. Nor is this lack of certainty necessarily a bad thing; a fallibilist attitude to one’s own beliefs has much to recommend it. Another difficulty for the full-belief view arises in connection with practical reasoning. On Bayesian principles, to assign a probability of 1 to a proposition is to cease to contemplate the possibility that it is false and, consequently, to ignore the undesirability of any outcome contingent upon its falsity. One consequence of this is that if one is certain of something, then one should be prepared, on pain of irrationality, to bet everything one has on its truth for no return at all. For one will simply discount the possibility of losing the bet. . . . Yet we can believe something, in the everyday sense, without being prepared to stake everything, or even very much, on its truth. (I
would bet something, but not a great deal, on the truth of my belief about my grandmother’s birth date). So flat-out belief is not the same thing as maximum probability.\footnote{Frankish continues:}

\[(\text{UNITY})\] is problematic for obvious reasons: it threatens robust skepticism about belief (many of the things we thought we believed, we don’t) and if the view is true then fallibilism about justified belief is false. \footnote{Fantl and McGrath 2009, 134}

There are, I think, two main versions of the worry that (UNITY) sets too strict a standard for outright belief (with some overlap between the two): the betting worry and the certainty worry.

The betting worry starts from some sort of connection between partial beliefs and betting behaviour, or betting dispositions.\footnote{Some connection of this sort is often given as an analysis, or interpretation, of the subjective probabilities found in the orthodox Bayesian formalism; however, see Eriksson and Hájek (2007) for a convincing argument that no such analysis can be satisfactory.} Roughly, one believes that $p$ to degree $x$ if and/or only if one is inclined to regard as fair a bet where one would win $S \cdot (1 - x)$ if $p$ is true and lose $S \cdot x$ if $p$ is false, where $S$ is some appropriate real number representing a stake. The idea is that, using $x$ as the probability that $p$ is true, the expected value of such a bet would be zero. On this picture, believing $p$ to degree 1 would mean regarding as fair a bet where one would stand to win nothing or nearly nothing should $p$ turn out to be true, but would lose some substantial amount

\footnote{There are several versions of the connection between credences and betting behaviour/dispositions attested in the literature. For my purposes, the differences among them—whether the property of interest is betting behaviour, or betting dispositions, or dispositions to regard certain bets as fair, or indifference between sides of a bet; whether betting behaviour/dispositions/etc. are supposed to provide an analysis of credence, or a way of measuring credence, or some other kind of connection—make no difference. In all versions, the problem with credence 1 is similar.}
should \( p \) turn out to be false. If it is part of the picture that \( S \) can take very large values, it is particularly unintuitive that ordinary people would regard bets at such long odds as fair, when \( p \) is some commonplace contingent proposition. Few people would regard as fair (in the sense of not favouring either side) a bet where they stand to gain a penny or lose their home on the proposition that Richard Nixon was once president of the United States, but surely most people should count as believing this proposition outright.

Given (CONTEXT), the betting worry looks much less plausible. The key insight is that offering a bet means changing the context. When the practical importance of \( p \) changes, as it must when a bet is offered, the space of salient alternatives to \( p \) may also change. When I am offered a bet on \( p \) at very long odds and/or very high stakes, I am likely to worry more about whether \( p \) is true after all, and consider a wider range of alternative possibilities. For example, were I to be offered a bet where I would win a dollar if Barack Obama is the current president of the United States or lose my home if he is not, I will think something like the following to myself: “Am I sure that I’ll win this bet? Is there something I might have been overlooking? I don’t want to lose all that money, and I don’t stand to gain much in comparison.” I will probably think of scenarios in which Obama is no longer president which I was previously ruling out or ignoring (maybe he resigned just a few minutes ago and the news has yet to reach me); I will be more careful about ruling those possibilities out, and may decide not to rule all of them out after all. If I do not rule out all unignored counterpossibilities to the proposition that Obama is president, then I will no longer give that proposition credence 1. I will also no longer believe that proposition, according to (UNITY).

I think this fits the standard intuitions about these extreme betting cases: when one stands to lose a lot if \( p \) is false, one is less likely not only to act on \( p \) but to assert sincerely that \( p \). Having had the thoughts above, I am still likely to be comfortable saying things like “Obama is probably president,” or “I’m pretty sure that Obama is president,” but I am not likely to say flat-out, “Obama is president.” One is especially likely to say the former sort of thing—that \( p \) is probable—if one is criticized for refusing a bet on \( p \): “I
thought you said that $p$—how can you turn down this bet?”—“Well, I really only think it’s very likely that $p$, but it’s not *that* likely.” All of this suggests that someone rejecting an extreme bet on $p$ does not believe $p$ while rejecting the bet.

(Of course, it’s also likely that I will reject the bet simply on the grounds that I do not want to risk my home at all, no matter how sure I am that the relevant proposition is true. But this is just to say that the betting worry about (UNITY) has somewhat less intuitive appeal than it might seem at first.)

Here is my response to the betting worry again, in a bit more detail. Assume (CONTEXT) is true. Suppose I have $Cr_1(p) = 1$, where $Cr_1(\cdot)$ is my probabilistically coherent credence function. Now suppose I am offered a bet according to which I would gain $1$ if $p$, or lose my home if $\neg p$. The expected value of this bet, according to $Cr_1(\cdot)$, is $1$, no matter what utility I assign to losing my home: for the expected value of the bet is given by $Cr_1(p) \cdot 1 + Cr_1(\neg p) \cdot u(\text{nohome})$, where $u(\text{nohome})$ is the (large negative) utility I assign to losing my home; since $Cr_1(p) = 1$ and $Cr_1(\neg p) = 0$, the expected value of the bet is $1 \cdot 1 + 0 \cdot u(\text{nohome}) = 1$. Therefore, I should accept the bet, since doing so has positive expected value.

But being offered this bet changes the practical stakes of $p$ for me. Before the bet was offered, I had no reason to think the falsity of $p$ might lead to my losing my home (or at least, this is the case for most $p$). The dramatically increased practical importance of $p$ can lead me to take erstwhile-ignored possibilities seriously; I may be more careful about what I am willing to rule out. Thus, my credences will no longer be given by the function $Cr_1(\cdot)$, but by a new function $Cr_2(\cdot)$. The fact that in the new context, where my credences are given by $Cr_2(\cdot)$, I take seriously some possibilities I ruled out in the old context, where my credences were given by $Cr_1$, means that there are some propositions $q$ such that $Cr_2(q) > Cr_1(q) = 0$. If there is such a proposition $q$ which I take to entail $\neg p$, and $Cr_2(\cdot)$ is probabilistically coherent, then it will be the case that $Cr_2(p) < Cr_1(p) = 1$. If $Cr_2(p)$ is low enough, then the bet under consideration will have negative expected value according to my revised credences, and so I will not accept it.
The certainty worry is vaguer than the betting worry, but less closely tied to problematic views about partial belief and betting behaviour/dispositions. Here, the trouble is that believing $p$ to degree 1 is interpreted as being certain that $p$—but sometimes people believe things without being certain. I think this worry is misguided: while I agree that mere belief does not require certainty, I do not think that certainty should be identified with belief to degree 1. A fully satisfying response to the certainty worry along these lines would require a more precise interpretation of just what it does mean to have degree of belief 1 than I am prepared to give here, but some remarks on certainty will give the flavour of a more complete response.

For one thing, being certain that $p$ seems to require some kind of stability of opinion, so to speak. If one is certain that $p$, then it must be difficult to move one to abandon belief in $p$. For example, part of what it means for me to be certain that the Earth is not flat, or that there are no genuine psychics, is that it would be very hard to bring me to disbelieve these propositions. It would take more, stronger, and better presented evidence to make me doubt that I was not nicknamed “Speedy” as a child, or that there are no genuine psychics, is that it would be very hard to bring me to disbelieve these propositions. It would take more, stronger, and better presented evidence to make me doubt that I was not nicknamed “Speedy” as a child (either believing that I did have that nickname after all, or withholding belief in either proposition) than it would to convince me that, say, my mother was nicknamed “Speedy” as a child. I believe that my mother had no such nickname, but all it would take to make me believe otherwise would be reliable-seeming testimony from one or two people who knew her as a child; this would not suffice to make me even doubt that I ever answered to that nickname. The difference is that while I believe both propositions, I am only certain of one. On the standard Bayesian formalism, degree of belief 1 is maximally stable for rational agents: as we saw in §4.2, if one has $Cr_1(p) = 1$, then there is no way to update by conditionalization to get $Cr_2(p) < 1$ (but see §4.6). Thus, at least for rational agents, if one ever gives a proposition credence 1, one can never

---

62I lean toward the approach to interpreting degrees of belief advocated in Eriksson and Hájek (2007), according to which we take degrees of belief as a primitive concept, and get at its meaning via the “platitudes” we take to be true about it. Thus, I have not much to say about what it means for one to have $Cr(p) = 1$ beyond giving a formal system for working with the function $Cr(\cdot)$. 

109
give it a lower credence; if either the threshold view of belief or (UNITY) is correct, then one will never withhold outright belief, either. But this need not be so if (CONTEXT) is true. For according to (CONTEXT), an agent’s degrees of belief vary with features of context other than her evidence. In particular, on the modified Bayesian formalism I propose here, credence functions are not always obtained by conditionalization on previous credence functions. It is possible to have $Cr_1(p) = 1$ and $Cr_2(p) < 1$, as we will see in §4.5 Therefore, if (CONTEXT) is true, it should be possible to believe $p$ to degree 1 without being certain that $p$, for certainty requires a stability of opinion that $Cr(p) = 1$ does not.

Thus, we see that neither the betting worry nor the certainty worry give good reason to reject (UNITY), provided that (CONTEXT) is true. I know of no other reason to reject (UNITY).

### 4.5 New Formalism

I have argued that (UNITY) solves a number of problems faced by the threshold view, and that (CONTEXT) makes (UNITY) plausible. Now I will present a formal framework for representing credences on which (CONTEXT) is true. The idea behind the formalism is to take the same approach to the representation of context, and of change of belief (or credence) across context, as we took in chapter 2. In §4.5.1, I will argue that the formalism described here generalizes the formalism of doxastic states described in chapter 2, given (UNITY).

Let a credal state $S$ for a set of propositions $P$ be a quadruple $(U_S, C_S, \|\|_S, Cr_S)$, where $U_S$ is a universe of points, $C_S$ is a set of subsets of $U_S$ (contexts), $\|\|_S : P \to U_S$ is a valuation function, $Cr_S : D_S \to [0, 1]$ is a conditional credence function (the agent’s “global” credence function), whose domain

---

63Although $Cr$ is a two-place function giving conditional (global) credences, we can use it to get unconditional credences by conditioning on a proposition the agent regards as logically true. That is, if $\|p\| = U$, then $Cr(\cdot|p)$ gives the agent’s unconditional (global) credences. We will see the reason for taking conditional rather than unconditional credences as primitive in §4.5.1.
$D \subseteq \mathcal{P}(U) \times \mathcal{P}(U)$ is as follows. Let $P' \subseteq \mathcal{P}(U)$ be the set $\{\|p\| | p \in P\}$ of the valuations of the propositions in $P$. Then $D$ contains at least the sets $U \times C$ and $P' \times C$. We write the value of $Cr$ on an ordered pair $(x, y) \in D$ as $Cr(x|y)$ if $y \neq U$, and call this value the agent’s credence (or degree of belief) in $x$, conditional on $y$ (or given $y$). (Note that, despite the notation, we do not stipulate that $Cr(x|y) = \frac{Cr(x \cap y)}{Cr(y)}$ when $Cr(y) > 0$. This is, in part, because we are not dealing only with ideally rational agents.) When $y = U$, we write $Cr(x)$, and call this value the agent’s (unconditional) credence in $x$. So, our constraint on $D$ amounts to ensuring that $Cr(\{x\}|C)$, and $Cr(\|p\| | C)$ are defined, for $x \in U$, $C \in C$, and $p \in P$. $D$ will usually be larger than required, with, e.g., $Cr^C(\|p\| | \|q\|)$ often being defined for $p, q \in P$ and $C \in C$.\footnote{As usual, the subscripts on $U, C, \|\cdot\|, D,$ and $Cr$ will be suppressed where there is no ambiguity.}

Sometimes, for convenience, we will write as if $Cr$ took propositions, rather than subsets of $U$, as arguments. In general, for $p, q \in P$, “$Cr(p|q)$” is to be understood as “$Cr(\|p\| | \|q\|)$”, and likewise for expressions with a name of a proposition in one argument-place and a name of a set in the other. Similarly, it will sometimes be convenient to have a shorthand for the proposition that possibility $x$ is actual. So, for every point $x \in U$, let $p_x$ be the proposition that $x$ is actual; that is, $\|p_x\| = \{x\}$. Call such a proposition the characteristic proposition for the point $x$. Likewise, for every context $C \in C$, let $p_C$ be the proposition that one of the members of $C$ is actual; that is, $\|p_C\| = C$. Call such a proposition the characteristic proposition for the context $C$. We can write a context’s characteristic proposition as the disjunction of the characteristic propositions of the members of that context: $p_C$ is equivalent to $\bigvee_{x \in C} p_x$.

If $C \in C$ is the current context, then the agent’s current degree of belief that $p \in P$ is given by $Cr(p|p_C)$; in other words, the agent’s degree of belief
belief that \( p \) in context \( C \) is given by the global credence in \( p \) conditional on the proposition that one of the possibilities \( x \in C \) is actual. We can thus define a function \( Cr^C : D \cap [\mathcal{P}(U) \times \mathcal{P}(C)] \to [0,1] \) giving the agent’s credences in context \( C \): let \( Cr^C(X|Y) = Cr(X|Y \cap C) \) for all sets \( X, Y \) with \((X, Y \cap C) \in D \cap [\mathcal{P}(U) \times \mathcal{P}(C)]\). Call \( Cr^C \) a “local” credence function. As usual, we will write \( Cr^C(\cdot) \) as shorthand for \( Cr^C(\cdot|p_C) \), for the agent’s unconditional local credences.

There is a point of interpretation to make here. In general, \( Cr(X|Y) \) or \( Cr^C(X|Y) \) is to be read as a conditional credence, the agent’s (global or local) credence that \( X \) given \( Y \). But there at least two ways that \( Cr(X|Y) \) could be read: speaking imprecisely, it could be how likely the agent now thinks \( X \) would be if she supposes \( Y \) were true, or it could be how likely the agent would think \( X \) is in case \( Y \) were known to be true. The former sort of reading has to do with what the agent thinks now; the latter with how her credences would change were \( Y \) true. In the case of \( Cr(X|C) \), i.e., those global conditional credences which define local unconditional credences, we must take the second sort of interpretation: \( Cr(X|C) \) is the credence one would have in \( X \) were it the case that \( C \), or in other words, \( Cr^C(X) \) is the credence one would have in \( X \) in the context \( C \). For all other conditional credences, including local conditional credences, either sort of interpretation is compatible with the formalism.

### 4.5.1 Credal States and Doxastic States

I claim that the models for (degrees of) belief introduced in this chapter generalize the models introduced in chapter 2, given (UNITY). Theorems 4.1 and 4.2 give the sense in which this is true, but first we need some definitions. (Recall that \textit{doxastic states} were the formal models introduced in chapter 2 for dealing with outright belief, and \textit{credal states} are the formal models introduced in this chapter for dealing with degrees of belief.)

**Definition 4.1.** A credal state \( S = (U_S, C_S, \|\|_S, Cr_S) \) and a doxastic state \( T = (U_T, C_T, \|\|_T, \preceq_T) \) agree, given (UNITY), just in case: \( S \) is a credal state for, and \( T \) a doxastic state for, the same set of propositions \( P; U_S = U_T \),
\( \mathcal{C}_S = \mathcal{C}_T \), and \( \| \cdot \|_S = \| \cdot \|_T \); and for any \( C \in \mathcal{C} \) and \( p \in P \), \( B_C \models p \) if and only if \( Cr_C(p) = 1 \).

The idea behind Definition 4.1 should be clear: a credal state and a doxastic state agree just in case in each context, the credal state assigns local credence 1 to exactly those propositions which doxastic state counts as believed.

**Definition 4.2.** Two doxastic states \( T_1 = (U_1, \mathcal{C}_1, \| \cdot \|_1, \preceq_1) \) and \( T_2 = (U_2, \mathcal{C}_2, \| \cdot \|_2, \preceq_2) \) are equivalent, and we write \( T_1 \sim T_2 \), just in case: \( U_1 = U_2, \mathcal{C}_1 = \mathcal{C}_2 \), and \( \| \cdot \|_1 = \| \cdot \|_2 \); and for any \( p \in P \) and \( C \in \mathcal{C} \), \( B^1_C \models p \) if and only if \( B^2_C \models p \), where where \( B^1_C \) is the set of \( \preceq_1 \)-minimal members of \( \mathcal{C} \), and \( B^2_C \) is the set of \( \preceq_2 \)-minimal members of \( \mathcal{C} \) for any context \( C \in \mathcal{C} \).

The idea behind Definition 4.2 is that if \( T_1 \sim T_2 \), then they agree on what the agent believes in every context.

**Definition 4.3.** Two credal states \( S_1 = (U_1, \mathcal{C}_1, \| \cdot \|_1, Cr_1) \) and \( S_2 = (U_2, \mathcal{C}_2, \| \cdot \|_2, Cr_2) \) are equivalent, and we write \( S_1 \sim S_2 \), just in case: \( U_1 = U_2, \mathcal{C}_1 = \mathcal{C}_2 \), and \( \| \cdot \|_1 = \| \cdot \|_2 \); and for all \( C \in \mathcal{C} \) and \( p, q \in P \cup \{ U \} \), \( Cr^C_1(p|q) \) is defined if and only if \( Cr^C_2(p|q) \) is defined, and \( Cr^C_1(p|q) = Cr^C_2(p|q) \) wherever they are defined.

As with equivalence of doxastic states, the equivalence of two credal states means that they accurately model exactly the same agents: they assign the same local credences to every proposition in \( P \).

As well as the above definitions, we need some constraints on the global credence function \( Cr \). So far, I have put no constraints on \( Cr \), but in order to derive the main result of this section—that credal states are a generalization of doxastic states—we need some constraints on \( Cr \) to match the constraints we put on the ordering relation \( \preceq \) in chapter 2. Putting constraints on \( Cr \) requires care, though, for we do not want to reduce the expressive power of the models for degrees of belief: we do not want it to turn out that the formal system is unable to represent genuinely possible partial belief states. In particular, that some constraint is met by all rational agents is
not sufficient reason for requiring that all agents meet that constraint. I will therefore begin by stating some formal constraints on $Cr$, and arguing that they do not render our formal system unable to represent any genuinely possible partial belief states.

There are three main constraints we put on $\preceq$: transitivity, reflexivity, and $\mathcal{C}$-well-foundedness. I will introduce two constraints on $Cr$ which will correspond to transitivity and $\mathcal{C}$-well-foundedness, respectively. We will see how this correspondence goes in the course of the proof of Theorem 4.1, and we will then also see why no constraint need be introduced to correspond to reflexivity.

**Transitivity Constraint** For any $x, y, z \in U$ and $C, D, E \in \mathcal{C}$, if $x \in C \cap E$, $y \in C \cap D$, $z \in D \cap E$, $Cr^C(\{x\}) = 0 < Cr^C(\{y\})$ and $Cr^D(\{y\}) = 0 < Cr^D(\{z\})$, then $Cr^E(\{x\}) = 0$.

According to this constraint, when we have a situation like that of Figure 4.1, the local credence functions $Cr^C$, $Cr^D$, and $Cr^E$ cannot induce a “cycle” on the points $x, y, z$. It helps here to think of $Cr^C(\{x\}) = 0$ as meaning that the possibility $x$ is ruled out in the context $C$. Then the Transitivity Constraint says that if there is a context where $x$ is ruled out and $y$ is not, and there is a context where $y$ is ruled out and $z$ is not, then in any context where both $x$ and $z$ are present, $x$ must be ruled out. (Cf. Figure 2.3 to see why I call the prohibited situation a cycle.)

**$\mathcal{C}$-Well-Foundedness Constraint** For all $C \in \mathcal{C}$, there is at least one $x \in C$ such that $Cr^C(\{x\}) > 0$.
I submit that it is unrealistic to suppose that there might be a context in which an agent thinks that there are no serious possibilities. For this constraint to be violated, there would have to be a context in which the agent gives every possibility credence zero, which is tantamount to taking no possibility seriously, or to ruling everything out. I see no reason why such cases should be allowed.

We need two further constraints on $Cr$ to get our results. In a sense, these are the constraints that do the work of getting the pairs of doxastic and credal states in Theorems 4.1 and 4.2 to agree.

**No Doubts (Nec)** For any $p \in P$ and $C \in C$, if $Cr^C(p) = 1$, then there is no $x \in C$ such that $x \not\models p$ and $Cr^C(\{x\}) > 0$.

**No Doubts (Suff)** For any $p \in P$ and $C \in C$, if $x \models p$ for every $x \in C$ such that $Cr^C(\{x\}) > 0$, then $Cr^C(p) = 1$.

These two constraints say that having no serious doubts that $p$ is necessary and sufficient for having local credence 1 that $p$. The idea is that for a point $x \in C$ to count as a serious doubt it must be such that $x \not\models p$ (i.e., it constitutes a doubt that $p$, or speaks against $p$), and it must be such that $Cr^C(p) > 0$ (i.e., it must be taken seriously).

Henceforth, by “credal state”, I shall mean a credal state as previously defined, which further satisfies the Transitivity Constraint, the $C$-Well-Foundedness Constraint, and No Doubts (Nec and Suff).

Now we are ready to show that credal states generalize doxastic states. That is accomplished by the following results.

**Theorem 4.1.** For any credal state $S$ there is a doxastic state $T$ such that $S$ and $T$ agree. $T$ is unique up to equivalence.

**Proof.** Let $S = (U_S, C_S, \| \|_S, Cr_S)$ be a credal state. We need to build a doxastic state $T$ that agrees with $S$, and show that it is unique up to equivalence. Begin by setting $U_T = U_T$, $C_T = C_T$, and $\| \|_S = \| \|_T$. Now we need to specify $\preceq$ such that $B_C \models p$ if and only if $Cr^C(p) = 1$ for all $C \in C$ and all $p \in P$. We will start by defining $B_C$ for each context: let
$B_C = \{ x \in C | Cr^C(\{x\}) > 0 \}$. Now let $x \prec^* y$ if and only if there is some context $C \in \mathcal{C}$ such that $x \in B_C$ and $y \notin B_C$; and $x \simeq^* y$ if and only if there is some context $C \in \mathcal{C}$ such that $x, y \in B_C$. Let $x \preceq^* y$ if and only if either $x \prec^* y$ or $x \simeq^* y$. Extend $\preceq^*$ to a new relation $\preceq^{**}$ by adding $(x, y)$ where $y \notin B_C$ for all $C \in \mathcal{C}$, and $x$ is any member of $U$. Thus, if a point does not belong to any belief set, then it is $\preceq^{**}$-maximal in $U$. Finally, let $\preceq$ be the transitive closure of $\preceq^{**}$—the smallest transitive relation containing $\preceq^{**}$.

Now we need to show that $\mathcal{T} = (U, \mathcal{C}, \|\cdot\|, \preceq)$ is a doxastic state. To show that $\mathcal{T}$ is a doxastic state, we must show that $\preceq$ is transitive, reflexive, and $\mathcal{C}$-well-founded. First, $\preceq$ is transitive by definition: it is the transitive closure of $\preceq^*$. Second, for any $x \in U$, if $x \in B_C$ for any $C \in \mathcal{C}$, then $x \simeq^* x$, and therefore $x \preceq x$. If, on the other hand, $x \notin B_C$ for all $C \in \mathcal{C}$, then $x \preceq^* x$, and so $x \preceq x$. Therefore, $\preceq$ is reflexive. Finally, $\preceq$ is $\mathcal{C}$-well-founded because the $\mathcal{C}$-Well-Foundedness Constraint on $Cr$ required that every context $C \in \mathcal{C}$ contain a non-empty subset all of whose members get non-zero credence. By our construction of $B_C$, these are just the members of $B_C$; it is easy to confirm that given the definition of $\preceq^*$, indeed $B_C$ is indeed the set of all $x \in C$ such that for all $y \in C$, $x \preceq^* y$. The move from $\preceq^*$ to $\preceq^{**}$ does not affect this fact, because $\preceq^{**}$ differs from $\preceq^*$ only by taking elements which are already not $\preceq^*$-minimal in any $C \in \mathcal{C}$, and making them $\preceq^{**}$-maximal on the whole universe $U$; therefore, the $B_C$ are also $\preceq^{**}$-minimal in $C$. And because $Cr$ meets the Transitivity Constraint, taking the transitive closure of $\preceq^{**}$ does not change the minimality of the $B_C$. That is, if there were a cycle $x \prec^{**} y \prec^{**} z \prec^{**} x$, then we would get $x \simeq y \simeq z$ when we take the transitive closure of $\preceq^{**}$. In such a case, if we had $x, y \in C \in \mathcal{C}$, and if $x$ were $\preceq^{**}$-minimal in $C$, then $y$ would not be $\preceq^{**}$-minimal in $C$, but would be $\preceq$-minimal in $C$. In order for there to be a cycle $x \prec^* y \prec^* z \prec^* x$, we would have to have contexts $C, D, E \in \mathcal{C}$ with $x \in C \cap E$, $y \in D \cap E$, $z \in D \cap E$, $Cr^C(\{x\}) = 0 < Cr^C(\{y\})$, $Cr^D(\{y\}) = 0 < Cr^D(\{z\})$, and $Cr^E(\{z\}) = 0 < Cr^E(\{x\})$. However, this arrangement of local credences is ruled out by the Transitivity Constraint, so there can be no cycles $x \prec^* y \prec^* z \prec^* x$. Furthermore, the move from $\preceq^*$ to $\preceq^{**}$ does not add any new cycles $x \prec^{**} y \prec^{**} z \prec^{**} x$, because if $v \prec^{**} w$
but \( v \not\prec^* w \), then we must have \( w \not\in B_C \) for all \( C \in \mathcal{C} \), so \( w \) must be \( \preceq^* \)-maximal in \( U \), so there can be no point \( u \) with \( w \prec^* u \). Thus, since there are no cycles \( x \prec^* y \prec^* z \prec^* x \), there can be no cycles \( x \prec^* y \prec^* z \prec^* x \). Therefore, the set \( B_C \) is indeed \( \preceq \)-minimal in \( C \) for all \( C \in \mathcal{C} \). Thus, \( \preceq \) is \( \mathcal{C} \)-well-founded. Therefore, \( \mathcal{T} \) is a doxastic state.

Next, we must show that \( \mathcal{T} \) agrees with \( \mathcal{S} \), given (UNITY). Suppose otherwise, for reductio. Then there is some context \( C \in \mathcal{C} \) and proposition \( p \in P \) such that (1) \( B_C \models p \) but \( Cr^C(p) < 1 \), or such that (2) \( B_C \not\models p \) but \( Cr^C(p) = 1 \). Suppose (1). Then, since \( B_C \models p \), \( x \models p \) for every point \( x \in C \) with \( Cr^C(\{x\}) > 0 \). But then by No Doubts (Suff), we must have \( Cr^C(p) = 1 \). Now suppose (2). Then there is some \( x \in B_C \) with \( x \not\models p \). Since \( x \in B_C \), \( Cr^C(\{x\}) > 0 \). But then by No Doubts (Nec), we must not have \( Cr^C(p) = 1 \). Thus, both (1) and (2) lead to a contradiction. Therefore, \( \mathcal{T} \) must agree with \( \mathcal{S} \).

Finally, we must show that \( \mathcal{T} \) is unique up to equivalence. Suppose there is another doxastic state \( \mathcal{T}' = (U', \mathcal{C}', \| \cdot \|', \preceq') \) which agrees with \( \mathcal{S} \). For every context \( C \in \mathcal{C} \), let \( B'_C \) be the set of \( \preceq' \)-minimal points members of \( C \). From the agreement of \( \mathcal{T}' \) with \( \mathcal{S} \), we have that: \( U' = U; \mathcal{C}' = \mathcal{C}; \| \cdot \|' = \| \cdot \|; \) and for every context \( C \in \mathcal{C} \) and every proposition \( p \in P \), \( B'_C \models p \) if and only if \( Cr^C(p) = 1 \). Then from the definition of \( \mathcal{T} \), it follows that \( B'_C \models p \) if and only if \( B_C \models p \). This entails that \( \mathcal{T}' \sim \mathcal{T} \).

**Theorem 4.2.** For any doxastic state \( \mathcal{T} \), there is at least one credal state \( \mathcal{S} \) such that \( \mathcal{S} \) and \( \mathcal{T} \) agree.

**Proof.** Let \( \mathcal{T} = (U_\mathcal{T}, \mathcal{C}_\mathcal{T}, \| \cdot \|_\mathcal{T}, \preceq_\mathcal{T}) \) be a doxastic state. We need to build a credal state \( \mathcal{S} \) that agrees with \( \mathcal{T} \). Begin by setting \( U_\mathcal{S} = U_\mathcal{T}, \mathcal{C}_\mathcal{S} = \mathcal{C}_\mathcal{T}, \) and \( \| \cdot \|_\mathcal{S} = \| \cdot \|_\mathcal{T} \). Now we need to specify \( Cr \) such that \( B_C \models p \) if and only if \( Cr^C(p) = 1 \) for all \( C \in \mathcal{C} \) and all \( p \in P \). For all \( x \in U, C \in \mathcal{C}, \) and \( p \in P \), we must define \( Cr^C(\{x\}) \), and \( Cr^C(p) \).

First, for each context \( C \in \mathcal{C} \), the members of \( C \setminus B_C \) are the possibilities ruled out in \( C \); so these must all get credence zero. Therefore, \( Cr^C(\{x\}) = 0 \) for each \( x \in C \setminus B_C \), and \( Cr^C(\{x\}) > 0 \) for all \( x \in B_C \). For all \( x \in B_C \), let \( Cr^C(\{x\}) = \alpha \), for some real \( \alpha \in (0, 1] \). Now we have defined \( Cr^C(\{x\}) \) for
all $x \in U$ and all $C \in \mathcal{C}$.

Next, for each context $C \in \mathcal{C}$ and proposition $p \in P$, we must have $Cr^C(p) = 1$ if and only if $B_C \models p$, in order to satisfy No Doubts (Nec and Suff). Suppose $B_C \not\models p$. Then $Cr^C(p) < 1$. If $x \not\models p$ for all $x \in B_C$, then set $Cr^C(p) = 0$. Otherwise, set $Cr^C(p) = \beta$, for some real $\beta \in (0, 1)$. Now we have defined $Cr^C(p)$ for all $p \in P$ and $C \in \mathcal{C}$.

Now we must show that $Cr$, so defined, satisfies all the constraints we have put forward so far. These are the Transitivity Constraint, the $\mathcal{C}$-Well-Foundedness Constraint, and No Doubts (Nec and Suff). First, for the Transitivity Constraint, suppose we have $x, y, z \in U$ and $C, D, E \in \mathcal{C}$ with $x \in C \cap E$, $y \in C \cap D$, and $z \in D \cap E$, $Cr^C(\{x\}) = 0 < Cr^C(\{y\})$, and $Cr^D(\{y\}) = 0 < Cr^D(\{z\})$. Then by construction of $Cr$, we must have $x \notin B_C$, $y \in B_C$, $y \notin B_D$, and $z \in B_D$. It follows, then, that $y \prec x$ and $z < y$. Therefore, $z < x$, by transitivity of $\preceq$. So $x \notin B_E$, because there is at least one member of $E$, namely $z$, with $x \not\models z$. Since $x \in E \setminus B_E$, we have $Cr^E(x) = 0$, by construction of $Cr$, which is just what the Transitivity Constraint requires. Second, for the $\mathcal{C}$-Well-Foundedness Constraint, we need to show that for every context $C \in \mathcal{C}$, there is at least one $x \in C$ with $Cr^C(\{x\}) > 0$. Because $\preceq$ is $\mathcal{C}$-well-founded, we know that every context $C \in \mathcal{C}$ has a non-empty $\preceq$-minimal subset $B_C$; and by construction of $Cr$, for every point $x \in B_C$, $Cr^C(\{x\}) = \alpha > 0$. Finally, No Doubts (Nec and Suff): we must show that for every $p \in P$ and $C \in \mathcal{C}$, $Cr^C(p) = 1$ if and only if $x \models p$ for all $x \in C$ with $Cr^C(\{x\}) > 0$. But the right-hand side of this biconditional is equivalent to $B_C \models p$, by construction of $Cr^C(\{x\})$, and we do indeed have $Cr^C(p) = 1$ if and only if $B_C \models p$, by construction of $Cr^C(p)$. Therefore, $\mathcal{S} = (U, \mathcal{C}, \|\|, Cr)$ is a credal state. And finally, the latter argument also shows that $\mathcal{S}$ agrees with $\mathcal{T}$, because agreement requires that for every $p \in P$ and $C \in \mathcal{C}$, $B_C \models p$ if and only if $Cr^C(p) = 1$. \qed

Theorem 4.1 shows that a credal state implicitly specifies a doxastic state. Theorem 4.2 shows that a doxastic state can always be embedded in a credal state which gives the same verdict about the agent’s outright beliefs. Thus, credal states carry strictly more information than doxastic states; one can
read off outright beliefs from a credal state, but not necessarily *vice versa*. The next result shows that one cannot, in general, read off degrees of belief from a doxastic state.

**Proposition 4.1.** Some doxastic states agree with multiple non-equivalent credal states.

*Proof.* Use the construction from the proof of Theorem 4.2. Provided there is at least one proposition $p \in P$ such that, for some $C \in \mathcal{C}$, $B_C \not\models p$ but there is an $x \in B_C$ such that $x \models p$, we can set $Cr^C(p) = \beta$ for any $\beta \in (0, 1)$. In particular, following the construction above taking $\beta = 0.2$ and taking $\beta = 0.8$ will produce two credal states $S_2$ and $S_2'$, with $S_1 \sim S_2$ because $Cr^C_1(p) \neq Cr^C_2(p)$. 

Thus, credal states (strictly) generalize doxastic states.

### 4.6 It’s Still OK to Be a Bayesian

In the final two sections of this chapter, I will give some discussion of the account of belief and credence described above. First, in this section, I argue that my credence-sensitivist view is compatible with a version of Bayesianism which improves on orthodoxy. In the next section, I will argue that my view has some further virtues.

What makes a theory of partial belief Bayesian? Well, two main things, as I see it: first, that ideally rational agents’ credences satisfy the probability calculus (probabilism); and, second, that ideally rational agents’ credences are updated by conditionalization. There are multiple plausible versions of probabilism compatible with the framework of §4.5. Minimally, probabilism should require that for each $C \in \mathcal{C}$, $Cr^C(\cdot)$ satisfies the probability axioms. But probabilists will generally want more than this, I think. For one thing, there is the constraint that whenever $Cr(x|y \cap z)$, $Cr(x \cap y|z)$, and $Cr(y|z)$ are all defined and $Cr(y|z) > 0$, we should have $Cr(x|y \cap z) = \frac{Cr(x \cap y|z)}{Cr(y|z)}$. This constraint does not follow from the probability axioms themselves, but see Hájek (2003, §9) for an argument that it is a good constraint for conditional...
probabilities. For the orthodox Bayesian, this constraint comes out as trivially true, by the definition of \( Cr(x|y) \); but unlike the orthodox Bayesian, I take conditional probabilities as primitive, rather than defining them in terms of unconditional probabilities. (Cf. note 53 on page 94 and note 65 on page 111.) We might slightly strengthen the constraint to demand that whenever \( Cr(y|z) \) is defined and positive, \( Cr(x|y \cap z) \) and \( Cr(x \cap y|z) \) should also be defined, for all \( z \subseteq U \) with \( y \cap z \neq \emptyset \). Finally, we might want the global credence function \( Cr(\cdot) \) itself to satisfy the probability axioms. (This last requirement would have to be interpreted as a requirement on how one’s credences should cohere across contexts, rather than as a synchronic requirement on one’s credences; one’s credences at any particular time are given by one of the local credence functions \( Cr_C \). We will return to this point below.)

Note, incidentally, that if the global credence function satisfies the probability axioms plus the above identity for conditional probabilities, then it follows that each of the local credence functions will also satisfy the probability axioms; but the converse implication is false. The important thing to note about all of these requirements is that they are all perfectly compatible with the credence-sensitivist framework.

The credence-sensitivist framework is also compatible with multiple versions of the Bayesian rule of updating by conditionalization. Minimally, we can require that \( Cr^C(p|q) \) be defined (i.e., that \( Cr(||p|| \cap ||q|| \cap C) \) is defined for all \( p, q \in P \)), and then say that when one learns in context \( C \) that \( q \), one’s new credence that \( p \) in context \( C \) should be \( Cr^C_{\text{new}}(p) = Cr^C_{\text{old}}(p|q) \). Things get more interesting when we look at how contexts must change both locally and globally, when one acquires new evidence. Locally, we might say that when one learns that \( q \), one’s context must shift in some specified way, say to the largest \( C' \subseteq C \) with \( \neg q \cap C' = \emptyset \). Globally, we might require moving from \( C \) to \( C' \), perhaps requiring that all contexts \( C' \in C' \) assign \( Cr^{C'}(q) = 1 \).

There is much to be explored in this direction, but that would be outside the scope of this dissertation. However, I think it is clear that the credence-sensitivist framework I have advanced is not incompatible with updating by conditionalization, in some form.

Aside from the question of whether our framework is compatible with
the two main components of Bayesianism, there is of course the question of whether the standard Dutch book arguments for each component still go through. In brief, there is good news and bad news: some Dutch book arguments still go through, but some do not. In particular, the diachronic Dutch book arguments used to justify updating by conditionalization do not work, because gaining new evidence can result in a change of context—new possibilities may come to be taken seriously. This will result in a change of the underlying space on which one’s local credence function is defined in what may not be a predictable (and hence Dutch-book-exploitable) way. On the other hand, synchronic Dutch book arguments may still go through. For example, it is still the case that if, in context $C$, one has $Cr^C(p) = 0.7$ and $Cr^C(\neg p) = 0.7$, then one will regard as favourable both of the following bets: (1) a bet that costs $1$ and pays $1.50$ if $p$; (2) a bet that costs $1$ and pays $1.50$ if $\neg p$. (Expected value of each bet: $(1.50)(0.7) - 1 = 0.05$.) But if one buys both bets, then one is guaranteed to pay $2$ but win back only $1.50—a sure loss of $0.50$. Here we must assume that $C$ is the context one would be in after the bets have been offered. In general, as has been pointed out, offering a bet is a way of bringing new possibilities into serious consideration. So synchronic Dutch book arguments can be expected to work, but only for the context in which a Dutch book has been offered; Dutch book arguments cannot show that one’s local credence function in other contexts ought to be probabilistic.

Of course, the so-called non-pragmatic or epistemic arguments for probabilism should also go through exactly as well on a sensitivist as on a non-sensitivist framework. See, e.g., Joyce (1998) and Leitgeb and Pettigrew (2010a,b).

So I would describe the formal framework of this chapter as broadly Bayesian, or at least Bayesian-friendly, although it certainly diverges from Bayesian orthodoxy. In fact, I think it is a clear improvement on orthodox

---

66 Or rather, they can be expected to work exactly as well as they work in a non-sensitivist framework. There are, of course, those who object to Dutch book arguments on other grounds, and their objections will not be any stronger or weaker on the present sensitivist framework.
Bayesianism. I will argue for this in two stages: first, I will show that the main victories of Bayesianism—for example, solving the ravens paradox, explaining the value of surprising or novel evidence and of diverse evidence, shedding light on the Quine-Duhem problem and Goodman’s new problem of induction (to follow the list from Earman 1992)—can be preserved on my account; second, I will argue that the differences between my framework and orthodoxy are more psychologically realistic, providing a response to some standard lines of criticizing Bayesianism for being too idealistic. Or, to look at it in a different way, one can see my account as describing agents who lie somewhere short of the extreme end of the Bayesian’s scale of rationality; that is, one can see my account as identifying a way that we mere humans fall short of the Bayesian ideal. Such a halfway-story helps us to put the Bayesian account to use for non-ideal agents.

The most celebrated Bayesian success stories depend on the Bayesian account of confirmation. Therefore, my strategy here will be to show that, on my view, we can hold on to the Bayesian account of confirmation, and thereby hold on to the Bayesian success stories. For a Bayesian, an evidence proposition \( e \) confirms (disconfirms, is confirmationally irrelevant to) a hypothesis \( h \) relative to background evidence \( k \) just in case \( \Pr(h|e \land k) \) is greater than (less than, equal to) \( \Pr(h|k) \). The probability function here is supposed to represent the credences of an ideally rational agent. Typical Bayesian explanatory successes involve showing that, given certain background information about the case to be explained, any agent whose credences are probabilistic and who updates them by conditionalization must have her credence in \( h \) raised (or lowered, or unchanged) by learning \( e \); or that her credence in \( h \) must be raised more by \( e_1 \) than by \( e_2 \), or that \( e \) raises her credence in \( h_1 \) more than in \( h_2 \), and so on. Now, given that my sensitivist framework is compatible with requiring that ideally rational agents have probabilistic credences and that they update their credences by con-

---

67I argue in Clarke (2010) that the Bayesian solution to the ravens paradox is not ultimately satisfying; but this is for reasons independent of sensitivism. The point here is that the Bayesian solution to the ravens paradox works exactly as well for the sensitivist as for the orthodox Bayesian.
ditionalization, it follows that the Bayesian account of confirmation is also compatible with my framework; and if this is the case, then the explanatory successes of Bayesian confirmation theory should carry over when the Bayesian moves to my sensitivist framework.

For example, consider the standard Bayesian solution to the ravens paradox. To be brief and to oversimplify, the Bayesian argues that since non-black things are much more numerous than ravens, discovering that a randomly-selected object is a non-black non-raven is much less improbable than discovering that the object is a black raven; accordingly, a rational agent’s credence in the hypothesis that all ravens are black—or, equivalently, that all non-black things are non-ravens—should be increased much more by finding a black raven than by finding a non-black non-raven. Here are the sorts of things that need to be true for the Bayesian solution to work (cf. Vranas 2004, §2 for the argument behind the Bayesian solution; here, “Ra” means that a is a raven, “Ba” that a is black, and “H” that all ravens are black.):

1. \( \text{Cr}(Ra)/\text{Cr}(\neg Ba) \) is minute. (Or, \( \text{Cr}(Ra) \ll \text{Cr}(\neg Ba) \).)
2. \( \text{Cr}(Ra|\neg Ba) > 0 \).
3. \( \text{Cr}(Ba|H) = \text{Cr}(Ba) \).
4. \( \text{Cr}(\cdot) \) satisfies the probability axioms.\(^6\)

The Bayesian wants to conclude that any agent whose relevant credences are rational and accurate will find that the hypothesis that all ravens are black is confirmed to a much higher degree by finding a black raven than by finding a non-black non-raven. 1 and 2 constitute the requirement that the agent have accurate credences: she should think that there are many more non-black things than there are ravens (1), but should not be already be ruling out the possibility of a non-black raven (2). 4 is a rationality requirement. 3 may

\(^6\)I write “\( \text{Cr}(\cdot) \)” rather than “\( \text{Cr}(\cdot|\cdot) \)” because orthodox Bayesians usually do not take conditional credence to be primitive, as I have, but rather take the equation \( \text{Cr}(p|q) = \frac{\text{Cr}(p \land q)}{\text{Cr}(q)} \), \( \text{Cr}(q) > 0 \), to define conditional credence. Cf. footnote 53 on page 94 and footnote 65 on page 111.
fall into either category; the point of Vranas (2004) is that no satisfactory argument for 3 has been offered.

There is no reason why this argument should not go through on my sensitivist framework, given that it goes through on the orthodox framework. So long as an agent’s credences are probabilistic, and meet the conditions 1–3, the conclusion of the standard Bayesian solution will hold. An agent who violates 1 or 2 misunderstands the situation of the ravens paradox; the standard Bayesian solution has nothing to say about such an agent. An agent who violates 4 is irrational by the Bayesian’s lights; the standard Bayesian solution has nothing to say about such an agent. An agent who violates 3 either misunderstands the situation or is irrational, depending on what the justification for 3 is supposed to be; the standard Bayesian solution has nothing to say about such an agent. It does not matter why the agent violates 1–4: even if it is because some shift in context has caused the agent to revise the possibility space underlying her credences, she still either misunderstands the situation or is irrational. In short: Bayesians can say what they want to say about all agents they want to say it about. Adding context-sensitivity to the picture saps none of the Bayesian’s strength.

Having determined that my sensitivist framework does not undermine the usual benefits of being a Bayesian, let us turn to the positive advantages of my framework over Bayesian orthodoxy. Perhaps most obviously, my framework makes room for the context sensitivity of partial belief, whereas the orthodox Bayesian must regard any such context sensitivity as irrational change of belief. It may be true that an ideal scientific reasoner would have context-insensitive beliefs and credences, but I have argued that humans do not. We cannot keep certain core platitudes about sincere assertion if our outright beliefs are context-insensitive; and taking credences to be context-sensitive makes better sense of the connection between outright and partial belief than any other view. Thus, at least, the sensitivist framework offered here allows us to identify a new way of approaching or falling short of the (orthodox) Bayesian ideal: our beliefs may be sensitive to contextual factors to which an ideal scientific reasoner’s beliefs would be insensitive. We can, then, strive to better approximate ideal reasoners by striving to keep a
relatively fixed space of possibilities in mind across various contexts. This, then, is clearly an improvement on the orthodox, non-sensitivist Bayesian framework.

Of course, I do not think it is at all obvious that we should regard a being with context-insensitive beliefs as an ideal for us humans to aspire to. After all, there are reasons why we reason as we do: our cognitive resources are limited, and they will not cease to be so. On the contrary, I think a theory of rationality ought to tell us how to respond to changes in context while making allowance for our limitations, i.e., recognizing that we cannot take seriously very many possibilities at once. But none of this is incompatible with the core of Bayesianism, since, as we have seen, probabilitism and updating by conditionalization can be required of a sensitivist agent. So a modified Bayesianism, adding some details about how credences should change in response to changes in context, should count as an improvement on orthodox Bayesianism.

Staying with the idea that a theory of rationality should recognize our limited resources, one of the criticisms levelled at Bayesianism by people like Cherniak (from whom we will hear shortly) is that probabilistic coherence requires phenomenal computational power. The credence-sensitivist framework advanced here helps alleviate this worry, though it may not be perfectly obvious why. The orthodox Bayesian agent has one credence function, updated only when the agent gains new evidence, whereas the credence-sensitivist agent has a different local credence function for each context, and must update her credences whenever there is a shift of context. Surely this takes more computational power than having only one credence function to keep track of! But I think it is a mistake to think that more credence functions means more cognitive load, so to speak. The orthodox Bayesian agent’s single credence function is much more unwieldy than the local credence functions of the sensitivist agent. This is because, generally, an or-

---

69 See the Appendix for an argument that, for reasons independent of context-sensitivity, Bayesians should take agents to have multiple credence functions. In particular, they should take agents to have a “global” prior function which is not abandoned on updating by conditionalization.
thodox Bayesian credence function must assign a value to every proposition about which an agent might have an opinion in some context or other, regardless of the agent’s current context. On the other hand, most propositions will be irrelevant to an agent’s credences in a particular context, and so need not be taken into consideration in determining the agent’s local credence function. (Formally, one can implement this feature by leaving $Cr^C(p)$ undefined for all $p$ irrelevant in $C$, by setting $Cr^C(p) = 0$ or $1$ for all irrelevant $p$, or by some other means; I do not commit myself here to any particular method of dealing with irrelevant propositions.) Thus, the sensitivist agent’s local credence functions each take much less cognitive power than the orthodox Bayesian’s credence functions.

Here, for instance, is Cherniak raising one particularly vivid reason to worry about the computational power required by the orthodox Bayesian:

Suppose that each line of the truth table for the conjunction of [a large set of] beliefs could be checked in the time a light ray takes to traverse the diameter of a proton, an appropriate “supercycle” time, and suppose that the computer was permitted to run for twenty billion years, the estimated time from the “big-bang” dawn of the universe to the present. A belief system containing only 138 logically independent propositions would overwhelm the time resources of this supermachine. Given the difficulties in individuating beliefs, it is not easy to estimate the number of atomic propositions in a typical human belief system, but 138 seems much too low.

(Cherniak 1986, 93)

(Cherniak is not explicitly targetting probabilism here, but I think his argument carries over, since probabilistic coherence requires assigning equal credence to equivalent propositions, and assigning lower credence to a proposition than to any of its logical consequences. Also, Cooper 1990 shows, albeit less colourfully that probabilistic updating is NP-hard.) This sort of worry is alleviated by our credence-sensitivist framework, as I argued in the previous paragraph. In any particular context, most propositions will be irrelevant,
and so can be regarded as equivalent without having to check lines of a truth table (or following any other non-trivial method). 138 might be far too low for the number of logically independent propositions an orthodox Bayesian credence function must check for equivalence, but it seems far too high for the number of logically independent propositions relevant to a given context.

But what about the sensitivist’s global credence function? This seems to have the same problem the orthodox Bayesian credence function is supposed to have: it must assign (a conditional) value to any proposition the agent has an opinion about in any context at all. The answer to this objection requires looking at how the various features of a credal state are to be interpreted.

A local credence function, obviously, represents the agent’s credences in the relevant context; these are things that are plausibly out there in the agent’s head, so to speak. The global credence function, on the other hand, need not be so interpreted. As I see it, all a global credence function tells us is what the agent’s local credence functions will be in various contexts; it need not correspond to anything in the agent’s head. Note, importantly, that since we took conditional credences as primitive, not as defined by the equation

\[ Cr(p|q) = \frac{Cr(p \land q)}{Cr(q)} \]

the fact that local credences are determined from the global credence function by the equation \( Cr^C(\cdot) = Cr(\cdot|C) \) does not entail that there is any particular coherence or coordination between the various local credence functions. To be sure, if the global credence function satisfies the probability axioms, and the now-familiar constraint on conditional credences, then there will have to be a great deal of coordination between local credence functions; but this can simply be taken to mean that the last of our proposed versions of probabilism is too strong, and should be regarded only as a particularly distant ideal. It is an advantage of our credence-sensitivist framework that so many shadings of probabilism are possible, making clearer the ways we can fall short of that ideal.

4.7 Some Further Virtues

In this section, I argue that my credence-sensitivist view has three nice features: it draws a satisfying connection between outright belief and action;
it answers the so-called Bayesian Challenge; and it justifies us in calling credences “degrees of belief”.

4.7.1 Belief and Action

Maher (1986) and Kaplan (1996, 103–5) point out that, if outright belief is compatible with credence less than 1, then the following plausible-sounding maxims must be false:

5. If one’s beliefs are rational, then it is rational to act on them. That is, if it is rational for one to believe that \( p \), then it is rational for one to act as if \( p \).

6. A rational person ought always to act in accordance with her beliefs, or as if her beliefs were true.

7. To believe that \( p \) is to be disposed to act as if \( p \).

Maher and Kaplan gloss “acting as if \( p \),” and the other related locutions above, as acting in a way that one believes would be optimal if \( p \) is true. 5 and 6 are common sense principles about rational action and belief; variations on 7 are endorsed by, e.g., Braithwaite (1932) and Marcus (1990, 140), as an analysis of belief.

Here is the reason none of these maxims would hold up. Suppose you believe that \( p \), but your credence in \( p \) is \( x < 1 \), and your credence in \( \neg p \) is \( (1 - x) \). Now if you must choose between an action \( A \) which is worth $0 whether \( p \) is true or false, and an action \( B \) which is worth $1 if \( p \) is true and \(-S(\frac{x+S}{1-x})\), for some positive number \( S \). The expected value of \( A \) is 0, and the expected value of \( B \) is \( x \cdot (1 - x)(\frac{x+S}{1-x}) = -S \); thus, \( A \) has a greater expected value than \( B \), and so if you are rational, you should do \( A \) rather than \( B \). But if \( p \) is true, then \( B \) is optimal: if \( p \) is true, then \( A \) is worth $0, and \( B \) is worth $1. So if you were to act according to your beliefs, i.e., as if \( p \) were true, then you would act irrationally. Furthermore, on some behaviouristic/dispositional interpretations of credence—e.g., if having \( Cr(p) = x \) means being disposed to accept any bet on \( p \) at odds \( x : 1 - x \) or
better—then it follows that you will not act according to your belief that \( p \) in this case, regardless of whether you are rational.

It should be clear that, since I endorse (UNITY), this problem does not arise for my account. If you believe that \( p \) in the context of the bet above, then you must have credence 1 that \( p \). Thus, some form of the maxims 5-7 are still available on my account. I say “some form of” those maxims because there is something importantly right in Maher and Kaplan’s arguments. Consider the following two variations on a principle like 5 and 6:

8. If \( S \) believes that \( p \) at \( t \), then at \( t \), \( S \) should act on \( p \).

9. If \( S \) believes that \( p \) at \( t \), and \( S \) has not revised her beliefs between \( t \) and \( t + \epsilon \), then \( S \) should act on \( p \) at \( t + \epsilon \).

I think 8 is right, but 9 is wrong. If \( S \) believes that \( p \) at \( t \), then \( S \) must have \( Cr(p) = 1 \) at \( t \); therefore, \( S \) should regard as optimal whatever action would be optimal if \( p \) is true. This is why 8 is right. But between times \( t \) and \( t + \epsilon \), there may have been some change in \( S \)’s context so that, despite not having changed or revised her credal state, she no longer has \( Cr(p) = 1 \). In particular, if she is offered a bet on \( p \) carrying the risk of a heavy loss if it turns out that \( \neg p \), \( S \) is likely to have her credence in \( p \) drop, due to thinking of new \( \neg p \)-possibilities which she was previously ignoring. Therefore, it might turn out that, at \( t + \epsilon \), \( S \) should not regard as optimal the action that would be optimal in case \( p \) is true, for exactly the reasons Maher and Kaplan point to.

Thus, credence sensitivism allows us to hold on to some of what is right about both the traditional doctrine that one ought to act on one’s beliefs and Maher and Kaplan’s arguments against that doctrine.

4.7.2 Answering the Bayesian Challenge

Kaplan (1996, 98ff) poses what he calls the “Bayesian Challenge”. Roughly, here it is: all we need for decision theory, and therefore all we need to explain rational action, is credence, not outright belief; so what is the point of talking
about outright belief at all? What work does outright belief do for us? What does it explain?

As Kaplan poses the Bayesian Challenge, it would be acceptable to answer that outright belief is reducible to credence, so talk about outright belief indirectly does the same work that talk about credence does. My account of belief and credence can easily answer this version of the Challenge: to believe outright that $p$ is to have $Cr(p) = 1$. But I think the Bayesian Challenge generalizes: even if belief is reducible to credence, we still have the question of whether we are ever better off talking about outright belief rather than credence. It would be very unsatisfying if belief-talk could always be replaced with credence-talk without losing anything. (Kaplan only presents the Bayesian Challenge after rejecting the possibility of reducing belief to credence—because he rejects the threshold view and (UNITY)—so it is not surprising that he gives the form of the Challenge that he does.) That is, we can see the sort of objection to the threshold view that I attributed to Stalnaker at the start of §4.3.2 as a generalization of the Bayesian Challenge. However, I have already addressed that objection in §4.3.2, so I will not discuss it any further here.

4.7.3 Degrees of What?

I have already mentioned, in note 60 on page 106 and note 62 on page 109, that I am impressed by the arguments in Eriksson and Hájek (2007) against analyses of credence in terms of betting behaviour (or anything like it), and in favour of taking credence as a primitive, unanalyzed term. Of course, one of the benefits of having an analysis of credence is that we would always know what we mean when we say that some agent has $Cr(p) = x$, or $Cr^C(p) = x$. On the other hand, if we take credence as primitive, then we must rely on general platitudes to tell us what it means for an agent to have $Cr(p) = x$. Now, some of the most central platitudes about credence have to do with the connection between credence and outright belief: in particular, credences are, in some sense, degrees of belief; to raise one’s credence that $p$ is to increase the degree to which one believes that $p$. I think this is one of the most central
platitudes about credence, in fact.

This is important for those of us who want to give an account of out-
right belief in terms of credence, for such an account may conflict with the
platitude. In particular, consider Kaplan’s “Assertion View” of belief (1996,
109–10): “You count as believing \( p \) just if, were your sole aim to assert the
truth (as it pertains to \( p \)), and your only options were to assert that \( p \), as-
sert that \( \neg p \) or make neither assertion, you would prefer to assert that \( p \).”
The Assertion View is supposed to meet the Bayesian Challenge because it
explains the importance of belief to a certain practice (“asserting hypotheses
in the context of inquiry”), but belief as defined by the Assertion View is not
reducible to credence. (As Kaplan puts it in stating the Bayesian Challenge,
“belief is not a state of confidence”.) This is, roughly, because the aim to
assert the truth involves weighing the goal of avoiding error against the goal
of comprehensiveness, and different people will weigh these goals differently;
indeed a single person can be expected to weigh the two goals differently with
respect to different propositions. I do not intend to give Kaplan’s Assertion
View a full treatment here; I certainly agree that highlighting the connec-
tion between belief and assertion is important (cf. chapter 2), the point I
want to make here is that the Assertion View seems to violate the platitude
quoted above. If belief that \( p \) is a disposition to prefer to assert that \( p \) rather
than assert \( \neg p \) or remain silent in a certain sort of context, it is not clear in
what sense credences are degrees of belief. To be sure, increasing credence
that \( p \) makes it increasingly likely that one believes that \( p \), but raising one’s
credence that \( p \) does not mean that one believes that \( p \) more than previously.

This is an advantage of both the threshold view and (UNITY) over views
like Kaplan’s.\(^{70}\) Given (UNITY) or the threshold view, the sense in which

\(^{70}\) The view described in Weatherson (2005)—on which, roughly, one be-
lieves that \( p \) just in case one’s preferences conditional on \( p \) are the same as
one’s unconditional preferences—is in the same boat as Kaplan’s, as I read
it. However, in two 2010 posts on his blog at http://tar.weatherson.org/
2010/04/03/pragmatics-and-justification/ and http://tar.weatherson.org/2010/
03/30/shorter-can-we-do-without-pragmatic-encroachment/, Weatherson describes
his (2005) as entailing a version of the threshold view on which the threshold for be-
lief varies across contexts. Such a view (also endorsed by Sturgeon 2008) would be in
the same boat as the threshold view and (UNITY), with respect to the argument of the
credences are degrees of belief, or partial beliefs, is obvious; and so leaving credence unanalyzed is relatively safe.
Conclusion

I have argued that both belief-sensitivism and credence-sensitivism are true. In chapter 1, I presented an ordinary scenario which is nicely accounted for by belief-sensitivism; in chapter 2, I argued that we need belief-sensitivism to keep the principle that an assertion of $p$ is sincere if and only if the assertor believes that $p$; in chapter 3, I presented a solution to the preface paradox which becomes available once we adopt belief-sensitivism, and argued that a solution of this sort is, intuitively, preferable to the more popular probabilistic solutions; and in chapter 4, I argued that credence-sensitivism nicely accounts for the relationship between partial and full belief, without running into the usual paradoxes. In chapters 2 and 4, I presented formal models of sensitivism belief and credence.

There are a number of interesting questions still to be pursued, which I will tackle in future work. One question that should be obvious after the discussion in chapter 1 of the connections between sensitivism and contextualism is whether contextualism about “believe” (and other words used to attribute belief) is plausible, given sensitivism. Also, I gave a preliminary sketch in chapter 1 of how the truth of sensitivism might bear on the debate over k-contextualism; now that I have argued for sensitivism, it makes sense to sort out in more detail what the implications of sensitivism are for contextualism. On the formal side of things, as noted at the end of chapter 2, it makes sense to make the model dynamic, so that it can represent belief revision. This also suggests giving a sensitivist-friendly account of rational belief, which would involve examining the question raised in §4.6 of how to understand the rule of updating by conditionalization within a sensitivist framework.

I think that adopting sensitivism opens up many rich questions to be ex-
plored, which should be of interest both to traditional epistemologists interested in contextualism and to formal epistemologists working on Bayesianism or on logical approaches to belief revision. There is a lot still to be done.
Bibliography


Weatherson, B.: 2005, ‘Can We Do Without Pragmatic Encroachment?’, *Philosophical Perspectives* 19, 417–43.


Appendix A

Giving a Probability Assignment a Memory

I am about to flip a coin. You believe that it will be flipped fairly. We can express your belief elegantly, and in good Bayesian fashion, by saying that your credences in Heads and Tails are as follows: $Cr_1(h) = Cr_1(t) = 0.5$.

I have now flipped the coin. It landed Heads. You see this, and update your credences. Your credences are now: $Cr_2(h) = Cr_1(h|h) = 1$; $Cr_2(t) = Cr_1(t|h) = 0$. But you still believe that the coin was flipped fairly. How are we to express your belief now?

It seems that this sort of situation should fit squarely in the Bayesian wheelhouse, so here is a challenge: find a way to represent your belief that the coin was flipped fairly without using any extra-Bayesian machinery, and in particular without resorting to the non-Bayesian language of binary belief. To be clear, I do not mean to claim that the apparatus of binary belief cannot be reconciled with the Bayesian apparatus of partial beliefs: on the contrary, I think binary belief can be reduced to partial belief. The point here is about the expressive power of the system of representing partial belief by quasi-probabilistic credence functions. It is a selling point of the Bayesian apparatus that there are some cases of belief that can be captured neatly without appealing to anything but credences. As we saw in the first paragraph of this note, the belief that a future coin flip will be fair is one of these cases. It seems that the belief that a past coin flip was fair should also be one of those cases, even when one has seen the result of the flip. Thus, I think that this is a challenge the Bayesian should take seriously.

It turns out that the challenge is not so easy to meet. Here are a few
inadequate responses.

We can’t say that \( Cr_2(h) = Cr_2(t) = 0.5 \), since this would contradict the results of conditionalization. On the usual model of updating by conditionalization, your old credence function \( Cr_1 \) is lost, replaced by \( Cr_2 \). This means we can’t use \( Cr_1(h) = Cr_1(t) = 0.5 \) to express anything about your beliefs after the coin is flipped, after updating your credences. And it won’t do to tell a story about some belief of yours about the way the coin is weighted, or what would have been the result of flipping the coin over and over again. The belief we’re trying to characterize is about the fairness of this specific act of flipping this coin. After all, even people who have no idea what a limiting frequency is can have beliefs about the fairness of coin flips. Nor can we say that \( Cr_2(h|s) = 0.5 \), where \( s \) is a proposition completely capturing the setup of the coin flip: since \( Cr_2(h) = 1 \), we must have \( Cr_2(h|s) = 1 \).

Perhaps some credence about objective probability (such as physical chance, or, more plausibly, epistemic probability) might be thought to do the trick: say, something like \( Cr_2(EP(h|s) = 0.5) = 1 \), where \( EP \) is an epistemic probability function based on your current evidence. The trouble with this characterization is that it is unrevisable by conditionalization: once we have \( Cr_2(p) = 1 \), we can never get to a revised credence function \( Cr_3 \) with \( Cr_3(p) < 1 \) via conditionalization. But you can, of course, learn that I did not flip the coin fairly after all.

So we must revise the suggestion here; perhaps instead we can characterize your belief by setting \( x \leq Cr_2(EP(h|s) = 0.5) < 1 \), for some suitably high \( x \). But this is only plausible if we say that one believes a proposition \( p \) if and only if one has credence at least \( x \) in \( p \). (Cf. Foley 1993.) In other words, the revised suggestion is just that we say you have an outright belief that the coin toss was fair. The revised suggestion, then, fails to meet our challenge.

But there is another way to answer the challenge, suggested by our first non-starter: don’t lose \( Cr_1 \) when you update. Keep that information. Skyrms (1983) describes a few ways of “giving a probability assignment a memory,” to use his phrase. Here is a version of one of Skyrms’s suggestions, applied to our case.
Rather than model agents as having a credence function which is replaced by a new one every time the agent adds some proposition to her evidence and updates accordingly, we model agents as having a single prior credence function $Cr$, plus a set $E$ of evidence propositions. (This does increase the load on the agent’s resources; but see the last section of Skyrms 1983, 160–61 for an argument that the increase might not be prohibitively large.) After I have flipped the coin and you have seen the result, your set of evidence propositions will include $h$, the proposition that the coin toss resulted in heads. We can express your current credences in Heads and Tails thus: $Cr(h|E) = 1; Cr(t|E) = 0$. Finally, we can express your current belief that the coin toss was fair as follows: $Cr(h|E \div h) = Cr(t|E \div h) = 0.5$, where “$\div$” denotes some appropriate subtraction or contraction operation. Belief revision theory (Hansson 2009) provides some candidates for “$\div$”; the idea is to modify the agent’s evidence set as little as possible while removing $h$ and anything that essentially depends on $h$.

But the point of this note is not to argue for one or another candidate for “$\div$”, or to argue that any particular way of giving credence functions a memory is theoretically satisfactory. Rather, what I hope to have done is provide a very basic motivation for giving credence functions a memory. The motivation here is to patch a gap in the expressive power of the Bayesian framework—if nothing else, that framework should be able to elegantly characterize simple beliefs about coin tosses! Unlike other discussions of giving credence functions a memory, I make no appeal to the problem of old evidence (Glymour 1980), which may after all require some other solution (see, e.g., Eells 1985 or Joyce 1999, 205ff for careful arguments that the problem of old evidence is actually several problems, some unanswerable by giving credence functions a memory). I hope to have given reason to think that, whatever the correct solution to the problem(s) of old evidence, we ought to seek a satisfactory way of giving credence functions a memory as a solution of what seems to be quite an independent problem.