

# Preface Writers Are Consistent

Roger Clarke

Submitted version (pre-peer-review);  
final version forthcoming in *Pacific Philosophical Quarterly*

## Abstract

The preface paradox does not show that it can be rational to have inconsistent beliefs, because preface writers do not have inconsistent beliefs. I argue, first, that a fully satisfactory solution to the preface paradox would have it that the preface writer's beliefs are consistent. The case here is on basic intuitive grounds, not the consequence of a theory of rationality or of belief. Second, I point out that there is an independently motivated theory of belief—sensitivism—which allows such a solution. I sketch a sensitivist account of the preface writer's (consistent) doxastic state.

## Introduction

A rational agent's beliefs must be consistent. So it seems, at least. The preface paradox (Makinson 1965) seems to show otherwise. Here is a version 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 she seems to have done nothing wrong. In fact, her case is quite mundane—modest prefaces are hardly unusual. And yet, on the

other hand, she violates what is widely regarded as a minimal constraint on rational belief: *don't be inconsistent*.<sup>1</sup>

The lesson many philosophers take from the preface paradox is that consistency is not always rationally required (e.g., Foley 1993; Christensen 2004; Sturgeon 2008, and indeed Makinson 1965). But in §3 I will argue that, on the contrary, Jennifer is not best represented as having inconsistent beliefs. It may be true that consistency is not a rational requirement, but the preface does not show it. There is, of course, still the work of explaining how Jennifer's beliefs could possibly be consistent given the statement of the paradox; in §4, I draw on the account of belief and assertion advanced in Clarke (under review), *sensitivism*, to sketch a representation of Jennifer's beliefs as consistent. The core insight is that there is a change of context between the preface and the body of the book; Jennifer's assertions in these different contexts commit her to qualitatively different beliefs, which turn out to be consistent.

## 1 Consistency

But first, let us pause to appreciate why the preface situation should seem paradoxical. That is, why would we think that rational belief must be consistent? It is easy to come out of an introductory logic class with the sense that inconsistent propositions are somehow distasteful things to be avoided, but one would expect we can provide a more substantive motivation for requiring consistency. I'll offer two alternative motivations, one shallow and one deep. (These two alternatives will resurface in §4.2.)

An inconsistent set of beliefs cannot all be true. Belief aims at truth, so an inconsistent set of beliefs is doomed to failure. Therefore, a rational agent will not have inconsistent beliefs. This, I claim, is a shallow motivation for requiring consistency. Rational agents have (at least) two competing aims in forming beliefs: paraphrasing William James, to believe the truth and to avoid error. An inconsistent believer cannot completely avoid error, but might enjoy the compensating benefit of believing many truths. And, as Foley (1993, ch. 4) argues, one may best satisfy these two competing aims by adopting a set of beliefs which cannot perfectly satisfy the aim of avoiding error. Here is an analogy (Foley 1993, pp. 163–4):

Imagine that you are given the opportunity to play the following game. There are ten cups on a table, numbered 1–10, and you know that nine of the ten cups each cover a pea. You are asked to predict of each cup whether or not it covers a pea. For each correct answer you receive \$1 and for each incorrect

---

<sup>1</sup>To be sure, consistency is still a condition on *ideal* rationality, since actual human beings generally have so many beliefs that we cannot realistically be expected to have ensured that they are all consistent. 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).

answer you pay \$1. What is the best strategy for you in the game? Bet “pea” on each cup.

In this game, the optimal strategy guarantees you will not have an ideal result. If you bet “pea” on each cup, you are guaranteed to be wrong about one cup. Nevertheless, your expected return from adopting this strategy is higher than from any other. Analogously, suppose you are considering ten propositions, such that you know exactly nine of those propositions are true, but you have no idea which is the false one. If the payoff for believing the truth and the penalty for believing the false combine analogously to the results of winning and losing bets in Foley’s pea game, then you should believe all ten propositions, despite knowing that you thereby cannot completely avoid error.

On the other hand, respect for deductive arguments gives us a deeper motivation for requiring consistency. A deductively valid argument shows that one cannot consistently believe the argument’s premises and the negation of its conclusion. If rational agents must have consistent beliefs, then we have an easy explanation for the normative force of deductively valid arguments: anyone who believes the premises but denies the conclusion of a valid argument is inconsistent and therefore irrational.<sup>2</sup> If rationality sometimes allows inconsistent beliefs, then rational agents must be able to ignore some valid arguments, believing their premises and denying their conclusions. We would then need an independent explanation of why other valid arguments are compelling.

To be sure, opponents of consistency as a rational requirement have responses to both the shallow and the deep motivations sketched above (see, e.g., Foley 1993, pp. 167–70). The point here is that the shallow motivation turns on a superficial feature of inconsistent beliefs. As Foley’s pea game shows, there are cases where one’s ideal strategy precludes an ideal result. Therefore, the fact that inconsistent beliefs must fall short of ideally satisfying the aim of avoiding error can only partially explain the *prima facie* rational requirement of consistency. On the other hand, there is an essential connection between inconsistency and deductive validity. It may be possible to find a response to this deeper motivation, explaining why valid arguments are compelling yet allowing rational inconsistencies; but this is more difficult than dismissing the shallow motivation.

I draw the contrast between these two motivations for a consistency requirement both to motivate the preface situation as genuinely paradoxical, and because we will see the two motivations come apart on the view of belief advanced in §4. There, I will argue for a solution on which we respect the deep motivation and dismiss the shallow one.

---

<sup>2</sup>If we take withholding belief—neither believing a proposition nor its negation—as a serious option, then the same considerations motivate a closure condition on rational belief as well as consistency, according to which rational agents must believe whatever is entailed by their beliefs. (Cf. Christensen 2004, §4.3 and Weatherson 2005, §5.) I ignore closure in the main text because it is less obviously relevant to the preface paradox. However, I see it as a point in favour of the present motivation for a consistency requirement that it also motivates a closure requirement. These two conditions seem closely related. The shallower error-avoidance motivation offered previously does not clearly motivate a closure requirement.

## 2 Solving the Preface Paradox

What would count as a solution to the preface paradox? 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 we are talking about.

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 " $p \wedge \neg p$ " instead of " $p \wedge q$ ". 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*,<sup>3</sup> not that she has expressed contradictory *propositions*.<sup>4</sup> 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 must show that, given what she says in her book and its preface, Jennifer is not best understood as having expressed contradictory beliefs. That is what I aim to do in this paper. In brief, my preferred 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. According to my sensitivist account of belief and assertion, this entails that the beliefs expressed by the preface disclaimer and by the body claims are not inconsistent. In broad terms, this is similar to what Evidentiary (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.

---

<sup>3</sup>We sometimes say that when someone has said that  $p$ , she has expressed the belief that  $p$ . Presumably, this is because of the plausible principle that one asserts sincerely that  $p$  iff one asserts that  $p$  while believing that  $p$ ; this principle will figure prominently in the solution I outline below.

<sup>4</sup>The word "expressed" is being used very loosely here; I mean to get at an intuitive distinction, not to make a technical point.

### 3 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.<sup>5</sup> Before I offer a response on Jennifer's behalf, note that the reason for this absurdity is not the form of

---

<sup>5</sup>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."

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  $p_1, \dots, p_n$ , from which  $q$  deductively follows, then one has asserted that  $q$  (perhaps on the additional condition that one recognizes the entailment of  $q$  by  $p_1, \dots, p_n$ ). 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  $p_1, \dots, p_n$ ? So you’re saying that  $q$ !” This is acceptable even when the link between  $p_1, \dots, p_n$  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. 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 there’s a chance I’m 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  $p_1 \dots p_n$  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-scheme is valid:

1. At least one of the beliefs I espouse in the body of my book is false.
2.  $p_1, \dots, p_n$  is an exhaustive list of the beliefs I espouse in the body of my book.<sup>6</sup>
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  $p_1, \dots, p_n$  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 Evidine and Roush read it, is of a different order than the body claims: if an author believes 1–3 above, and also believes  $p_1, \dots, p_n$ , then she thereby has inconsistent beliefs. So, in other words, one might worry that Evidine'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.<sup>7</sup>

The argument I offer in this section heads off that worry. Professor Socrates and Jennifer agree on the inventory of claims she has made, and they agree on what makes a

---

<sup>6</sup>The version of the preface that Evidine 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 Evidine 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  $p \wedge q$ ) 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.

<sup>7</sup>Christensen (2004, 37) makes roughly this point.

belief false. Nevertheless, Jennifer insists, reasonably, that she has not committed herself to the denial 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.

The point of the argument here turns on an important distinction. 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 justifiably, 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. To put the point more contentiously: rejecting consistency as a rational requirement is not enough to resolve the preface paradox.

## 4 A New Solution

It is to our advantage if we can take what Jennifer says seriously, but it is not obvious how to do so. A *sensitivist* account of belief along the lines of Clarke (2012, 2013, under review) lets us take her claims at face value. This section explains, informally, how that would work.

Here is what I mean by “taking what Jennifer says seriously.” We want a treatment of Jennifer's exchange with Professor Socrates according to which she believes what she



says, and what she says at the end of their dialogue is true. Thus, we need: that Jennifer believes  $p_1, \dots, p_n$ , the propositions she asserts in her talks, when she asserts them; that she believes that, probably, at least one of  $p_1, \dots, p_n$  is false, when she asserts this at the end of the dialogue; that she has not changed her mind or revised her beliefs in the meantime; and that this does not reveal an inconsistency in her beliefs. On the face of it, this seems to be an impossible task, but sensitivism gives us a way of accomplishing it. §4.1 sets out the basic sensitivist account of belief, and then §4.2 considers how to incorporate a consistency norm into the sensitivist framework. Finally, §4.3 argues that preface writers do not violate any well-motivated consistency norm, given sensitivism.

## 4.1 Sensitivism

Belief is context-sensitive. This is the sensitivist's central claim. To be more precise: whether one believes that  $p$  depends, in part, on the space of serious alternatives, and this space can change from context to context. In particular, changes in the practical importance of  $p$  and in the conversational salience of alternatives to  $p$  tend to change the space of serious alternatives to  $p$ . An agent may count as believing that  $p$  in one context but not another, without any change in her doxastic state between the two contexts. *Doxastic state* is a technical term here, meant to generalize the more familiar notion of a belief set. The thought is that if belief is context-sensitive, then to completely specify an agent's beliefs, we need more than just a list of the propositions she believes: we need such a list for each context she might find herself in.

It may be easiest to think of this brand of sensitivism as a “relevant alternatives” account of belief, somewhat analogous to relevant alternatives accounts of knowledge (e.g., Dretske 1970): if there is some possibility  $x$  where  $p$  is false, and  $x$  is a relevant alternative in one context but not the other, and one's doxastic state does not dispose one to rule  $x$  out, then one will count as believing that  $p$  only in the context where  $x$  is not relevant. (Naturally, what makes an alternative relevant, and what it means to rule a possibility out will be different on this account than on standard “relevant alternatives” accounts of knowledge, but the parallels should be clear.)

Why think sensitivism is right? A full defense of sensitivism would be outside the scope of this paper,<sup>8</sup> but a sensitivism-based solution to the preface paradox will seem *ad hoc* without giving some motivation for the view. So here is a brief summary of some selling points of sensitivism.

*Case Pairs* Sensitivism receives support from some of the same case pairs used to motivate epistemic contextualism, such as DeRose's Bank Cases (DeRose 1992). Contextualists think that the truth conditions of sentences like “ $S$  knows that  $p$ ” are sensitive to the speaker's context—in particular, to the practical importance of  $p$  for the speaker, and to the conversational salience (to the speaker) of alternatives to  $p$ . Sensitivists think that belief is sensitive to

---

<sup>8</sup>See instead Clarke (under review).

the same factors of the subject's context (i.e., *S*'s context). Contextualists sometimes use first personal case pairs to motivate their view: we have two cases where the protagonist says something like "I know that *p*;" the subject speaks truly in one case and falsely in the other; and the only changes across cases are in practical importance and salience of alternatives. Sensitivism can explain the difference in verdicts on the protagonist's utterance: if the subject believes that *p* in only one of the cases, then the subject knows that *p* in at most one of the cases, since knowledge entails belief. (See Clarke (2012, ch. 1) for a detailed argument.)

*Degrees of Belief* Clarke (2013) argues that, by extending sensitivism to degrees of belief as well as outright belief, we can resolve the notoriously difficult problem of giving a unified account of both notions without running into the usual paradoxes. In short: once we recognize that degrees of belief are context sensitive in the right way, the usual reasons for rejecting the identification of outright belief with degree of belief 1 evaporate; and the latter identification avoids the usual objections to views on which outright belief is compatible with degree of belief less than 1.

*Assertion* Clarke (under review) argues that we need sensitivism to hold on to the platitude that an assertion of *p* is sincere if and only if the speaker believes that *p*. This is because the sensitivist account of belief mirrors the widely accepted Stalnakerian account of assertion (Stalnaker 1973, 1974, 1978, 1998, 2002). Stalnaker characterizes the conversational context in which an assertion is made by a set of possible worlds, namely those compatible with the information in the *common ground* of that context. An assertion that *p* aims to eliminate from the context set all worlds where *p* is false—but this will mean eliminating different possibilities in different contexts, depending on what information is in the contexts' common ground. This shiftiness in the effect of an assertion that *p* requires a matching shiftiness in what belief that *p* amounts to, and sensitivism is what fits the bill.

The last point, about sincere assertion, will be particularly important for our treatment of the preface, since our only evidence about Jennifer's beliefs comes from her (sincere) assertions.

Having given some independent motivation for sensitivism, let's put some flesh on the skeletal version of the view we've seen so far. An agent's beliefs at a given time are characterized by a *doxastic state*, consisting of the following elements. We have a set of possible worlds (possible according to the agent, that is), an ordering on those worlds, and a collection of sets of possible worlds. The sets of possible worlds are our contexts: we characterize a context by the set of possibilities taken seriously in that context. The ordering gives us something like a plausibility relation on the possible worlds: if world *x* is ahead of world *y* in the ordering, then the agent is disposed to rule out *x* in favor

of  $y$  in any context in which both  $x$  and  $y$  occur. We assume that the ordering gives us a non-empty set of “most plausible” worlds for each context. Finally, we say that an agent believes that  $p$  in a context  $C$  just in case  $p$  is true at all the most plausible worlds in  $C$ . What can lead to a shift in context? Anything that would lead to a change in the possibilities taken seriously. Typical context-shifting maneuvers include explicitly mentioning some previously unconsidered possibility, and altering the practical importance of some well-chosen proposition. (If I have to stake my house on correctly naming the current president, I will be sure to rule out even unusual sources of error.)

This is enough detail to state the condition on sincere assertion defended in Clarke (under review). We borrow from Stalnaker the ideas that a conversational context is characterized by the set of possible worlds compatible with the information in the common ground (i.e., roughly, the information commonly believed by speaker and audience), and that an assertion typically aims to shrink this set by adding the asserted proposition to the common ground. Then we say that an assertion that  $p$  is sincere iff, where  $C$  is the set of possible worlds the assertor takes to be compatible with the common ground in the context of assertion,  $p$  is true at all the maximally plausible worlds in  $C$  (maximally plausible according to the speaker). This is meant to be a precisification of the platitude “One asserts sincerely iff one believes what one asserts.” We will apply this view of sincere assertion in §4.3 to see what Jennifer’s assertions tell us about her doxastic state.

## 4.2 Consistency for Sensitivists

But first, we must see what it would mean to apply a consistency norm to our sensitivist framework. Sensitivism is a view about belief, but the preface paradox concerns *rational* belief. We still need to add norms of rationality to the sensitivist framework before we can produce a sensitivist treatment of the paradox.

On the sensitivist picture of belief, it is not obvious how we should read acontextual claims that an agent has inconsistent beliefs. Of course, believing both  $p$  and  $\neg p$  (or some larger set of mutually inconsistent propositions) in one and the same context is every bit as bad on a sensitivist view as on a conventional view of belief—but what should we say about someone who believes  $p$  in one context and  $\neg p$  in another context, and never believes both at once? I claim that the latter condition does not violate any plausible consistency norm.

To see this, recall the deep and shallow motivations for the consistency norm on the ordinary, context-insensitive view of belief outlined in §1. The shallow motivation appealed to the fact that an inconsistent set of beliefs cannot all be true. The deep motivation, on the other hand, pointed to the normative force of valid arguments. Once we move to a sensitivist picture of belief, we can see that the shallow motivation might suggest an cross-contextual consistency requirement, but the deep motivation does not. That is, if we want to avoid guaranteeing that we sometimes believe false things, then we should avoid believing  $\neg p$  in any context, if there is a context where we believe  $p$ .

Otherwise, we are bound to be mistaken in one of the two contexts.

On the other hand, provided only that we never believe an inconsistent set of propositions in any single context, we will never find ourselves faced with a valid argument whose import we ignore. We expect arguments to have persuasive force when they are voiced or considered—that is, considered all together. If  $p$  and  $q$  entail  $r$ , then we want anyone who considers  $p$ ,  $q$ , and  $r$  all together not to believe  $p$ ,  $q$ , and  $\neg r$ . But we need not expect that anyone who has ever believed  $p$  and believed  $q$  would never, under any circumstances, believe  $\neg r$ . We face arguments all at once, in a single context. Therefore, respect for valid arguments' normative force requires only consistency within contexts, not consistency across contexts.

### 4.3 Preface Writers Are Consistent

Now consider what we can learn about Jennifer's beliefs from her sincere assertions. If we know that Jennifer sincerely asserted  $p$  in some context, all we know is that  $p$  is true in the worlds most plausible to Jennifer in that context. If she also sincerely asserted  $\neg p$  in some distinct context, we cannot infer, without further information, that there is any context where Jennifer believes a contradiction. This will be essential to my solution.

I claim that there is a significant change of context between the preface and the body of Jennifer's book. While writing the body (that is, while sincerely asserting the body claims) Jennifer means to rule out possibilities having to do with her evidence for particular claims. On the other hand, while writing the preface (while sincerely asserting the preface disclaimer) she means to rule out possibilities having to do with her fallibility in general. That is, in the preface but not in the body, she and her intended audience take into consideration questions about how likely it is that she has written what would probably be the first completely error-free history book of any length. Because the beliefs expressed in the body and in the preface do not involve ruling out the same counter-possibilities, they can be integrated into (derived from) an overall doxastic state which does not exhibit any odious sort of inconsistency.

This explains what is wrong with Professor Socrates's criticism. When Jennifer gave her lectures, she was speaking in contexts analogous to those in which she wrote the body of her book. When she makes claims in the course of her lectures, she is aiming to rule out possibilities having to do with, say, ways history could have been, but not having to do with her own fallibility. But Professor Socrates explicitly raises to conversational salience possibilities in which Jennifer has made some mistake. This changes the conversational context so that Jennifer cannot continue to make assertions about history without also ruling out possibilities about her own fallibility: so she could not, in her final response to Professor Socrates, repeat her claims from the lectures without making the kind of immodest claim to infallibility that he accuses her of having already made. However, she can make a disclaimer about her own fallibility without retracting what she claimed in the lecture. This is what she does in the response I offered on her behalf on page 6. In effect, she refuses to rule out the possibility that she is in error, and

so *a fortiori* refuses to rule out all possibilities in which her claims are at least partly false; but she also refuses to reverse her position on the possibilities ruled out during the talk.

It is natural to suppose that once Jennifer moves on to the next questioner, possibilities about her own fallibility will fade into the background again: they were raised to salience by Professor Socrates in order to make a specific objection; once that objection is considered defeated, those possibilities need not be considered any longer. Thus we see that on my account, contexts can shift quite fluidly. This is not to say that “anything goes”—we can give a more detailed account of just how and when contexts shift, in general; but to explain in full how this would work is the business of another paper.

Of course, the mere fact that Jennifer makes her mutually inconsistent assertions in distinct contexts does not, by itself, entail that there is no context where she believes them all. To show that would be to show too much—after all, there are surely some people with inconsistent beliefs. But to solve the preface paradox, it is enough to show that one can make the preface disclaimer without having inconsistent beliefs. Jennifer’s reply to Professor Socrates points to a way that one can do so: Jennifer asserts the preface disclaimer in a relatively stringent context, where certain possibilities (possibilities where her human fallibility has led to an error somewhere in the book) are attended to; these possibilities are ignored in the context where she asserts the body claims; but she regards these possibilities as more plausible than the ones attended to in the body contexts. As a consequence, she does not believe the body claims when she asserts the preface disclaimer, though she may believe that each body claim is very likely to be true. This is not an extraordinary state we ascribe to Jennifer, and it is not inconsistent.

Here is an objection. My solution has it that Jennifer fails to have inconsistent beliefs within any context because she is attending to different sorts of alternatives in the body contexts and in the preface context. But one might worry that this does not address the most pressing form of the preface paradox. As I have presented it, the paradox has to do with a certain mundane situation in which it seems that Jennifer has done nothing wrong, rationally speaking, and yet it also seems that she has violated a basic consistency norm. The solution I offer explains why we need not see Jennifer as having violated any plausible consistency norm. But now consider a slightly different version of the paradox. Forget about what situations are mundane, and forget how ordinary people like behave. Consider instead how they ought to behave. It might seem that if Jennifer considers, all at once and in the same context, her evidence for her individual body claims and her evidence about the prevalence of errors in books like hers, then she ought to believe each of the body claims and also that at least one of the body claims is mistaken. Thus, regardless of what preface writers typically do believe, they ought to believe inconsistently.

I find this form of the paradox less interesting than the version I have been addressing so far. The preface situation is only plausibly paradoxical—in either form—if the book contains a great number of claims. If Jennifer can consider all the contents of her book at once, not to mention the evidence for all her claims, she is not much like us humans. Plausibly, different norms of rationality apply to creatures with significantly

different rational capacities than ours. On the other hand, if Jennifer's book is short enough that a mere human could consider it all together in one context, it is less clear that she ought to believe inconsistently. Consider the relative rarity of preface-paradox-style disclaimers attached to works the length of a journal article.

Perhaps, though, the point of this version of the paradox is supposed to be that an ideally rational version of Jennifer should have inconsistent beliefs. I think even this is wrong: it is not the case that Jennifer ought to believe inconsistently if she were to consider the totality of her evidence about the body claims and about her own fallibility in a single context, because evidence of the latter sort undermines beliefs of the former sort. If Jennifer is thinking seriously about her own fallibility, then she ought not to believe any of the body claims outright, though she ought still to believe things in the neighborhood: where  $p$  is one of the body claims, she ought not believe that  $p$ , but that probably  $p$ , or that the  $p$  is most likely given her evidence.<sup>9,10</sup> To abstract away from concerns about how much information a human mind can handle at once, consider a preface-like situation where there are only two body claims,  $p$  and  $q$ . This will be a case where Jennifer has excellent (but defeasible) evidence for each of  $p$  and  $q$ , as well as excellent evidence that that she is mistaken about at least one of the two. (Suppose basketball teams A and B are to play against each other under NBA rules;  $p$  is the proposition that A will not win, and  $q$  is the proposition that B will not win. One expert handicapper tells Jennifer that  $p$ , and another tells her that  $q$ . Jennifer knows that draws are possible under NBA rules, but extremely unusual.) In this sort of situation, it is very clearly not the case that Jennifer ought to believe all three of  $p$ ,  $q$ , and  $\neg(p \wedge q)$ . Rather, she should believe things in the neighborhood of  $p$  and  $q$ , such as that her evidence indicates that  $p$  (and that  $q$ ). The same is true for an ideal agent who can simultaneously consider the evidence for a book's worth of claims.

Therefore, we can safely say that the story told above about how to reconcile Jennifer's beliefs about the preface disclaimer and the claims in the body of her book does its job as advertised: we have a new solution to the preface paradox.<sup>11</sup>

---

<sup>9</sup>We might also say that it is still appropriate for Jennifer to give  $p$  high credence, but credence short of 1. I argue elsewhere for identifying outright belief with credence 1, but that is outside the scope of this paper.

<sup>10</sup>Of course, it would still be appropriate for Jennifer to make the unhedged assertion that  $p$  in the body of her book, rather than that probably  $p$ ; this is because, whatever alternatives Jennifer might happen to be considering as she writes, she should expect her audience not to be taking possibilities about Jennifer's general fallibility as live ones in the conversation. That is, Jennifer should expect that the bare assertion that  $p$  will not be taken as aiming to rule out the possibility that Jennifer is fallible in the ways that typical academic authors are fallible. (On the other hand, if Jennifer were writing for a community of superhuman beings of the sort mentioned above, who can consider all the evidence for the claims in the body of Jennifer's book at once, the conventions and norms around assertions in academic books might be different.)

<sup>11</sup>Versions of this paper were presented at the 2011 Notre Dame-Northwestern Graduate Epistemology Conference, at Simon Fraser University, and at the University of Southampton; my thanks to all three audiences. Thanks also for helpful discussion and comments to Paul Bartha, David Christensen, Jonathan Jenkins Ichikawa, Carrie Ichikawa Jenkins, Masashi Kasaki, Adam Morton, Max Weiss, and especially to

## References

- Carroll, L.: 1895, 'What the Tortoise Said to Achilles', *Mind n.s.*, **4**, 278–80.
- Cherniak, C.: 1986, *Minimal Rationality*. MIT Press, Cambridge, MA.
- Christensen, D.: 2004, *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford, Oxford.
- Clarke, R.: 2012, *Belief in Context*, PhD dissertation, University of British Columbia.
- Clarke, R.: 2013, 'Belief Is Credence One (In Context)', *Philosophers' Imprint* **13**(11), 1–18.
- Clarke, R.: under review, 'Assertion, Belief, and Context'. Available at <http://www.rogerclarke.org/research>.
- DeRose, K.: 1992, 'Contextualism and Knowledge Attributions', *Philosophy and Phenomenological Research* **52**(4), 913–29.
- Dretske, F.: 1970, 'Epistemic Operators', *Journal of Philosophy* **67**, 1007–23.
- Evnine, S. J.: 1999, 'Believing Conjunctions', *Synthese* **118**, 201–27.
- Foley, R.: 1993, *Working Without a Net: A Study of Egocentric Epistemology*. Oxford, Oxford.
- Goldman, A.: 1986, *Epistemology and Cognition*. Cambridge, Cambridge.
- Hacking, I.: 1967, 'Slightly More Realistic Personal Probability', *Philosophy of Science* **34**, 311–25.
- Kitcher, P.: 1992, 'The Naturalists Return', *Philosophical Review* **101**, 53–114.
- Makinson, D. C.: 1965, 'The Paradox of the Preface', *Analysis* **25**(6), 205–7.
- Pollock, J.: 1995, *Cognitive Carpentry: A Blueprint For How To Build A Person*. MIT Press, Cambridge, MA.
- Pollock, J.: 2006, *Thinking about Acting: Logical Foundations for Rational Decision Making*. Oxford, Oxford.
- Roush, S.: 2010, 'Optimism about the Pessimistic Induction', in P. D. Magnus and J. Busch (eds.), *New Waves in Philosophy of Science*, pp. 29–58. Palgrave-MacMillan, New York.
- Stalnaker, R.: 1973, 'Presuppositions', *Journal of Philosophical Logic* **2**, 447–57.

---

Chris Stephens.

- Stalnaker, R.: 1974, 'Pragmatic Presuppositions', in M. K. Munitz and P. K. Unger (eds.), *Semantics and Philosophy*, pp. 197–214. New York University Press, New York.
- Stalnaker, R.: 1978, 'Assertion', *Syntax and Semantics* **9**, 315–32. Reprinted in Stalnaker (1999, 78–95). Page references are to the reprinted version.
- Stalnaker, R.: 1998, 'On the Representation of Context', *Journal of Logic, Language, and Information* **7**, 3–19.
- Stalnaker, R.: 1999, *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford, Oxford.
- Stalnaker, R.: 2002, 'Common Ground', *Linguistics and Philosophy* **25**, 701–21.
- Sturgeon, S.: 2008, 'Reason and the Grain of Belief', *Noûs* **42**(1), 139–65.
- Weatherson, B.: 2005, 'Can We Do Without Pragmatic Encroachment?', *Philosophical Perspectives* **19**(1), 417–43.