



Once you think you're wrong, you must be right: new versions of the preface paradox

John N. Williams¹ 

Received: 12 April 2017 / Accepted: 12 September 2018 / Published online: 19 September 2018
© Springer Nature B.V. 2018

Abstract

I argue that there are living and everyday case in which rationality requires you, as a non-idealized human thinker, to have inconsistent beliefs while recognizing the inconsistency. I defend my argument against classical and insightful objections by Doris Olin, as well as others. I consider three versions of the preface paradox as candidate cases, including Makinson's original version. None is free from objection. However, there is a fourth version, Modesty, that supposes that you believe that at least one of your beliefs (excluding this) is false. I argue that this version escapes all the objections that could trouble the other versions, including the objection that given certain closure principles for justified belief, justified inconsistent beliefs saddle you with a pair of justified beliefs that are in explicit contradiction (their contents being syntactic negations). I also argue more tentatively for the same verdict for Modesty*, a version that supposes that you believe that at least one of your beliefs (including this) is false. In that case, your belief guarantees that its content is true, in which case your beliefs are inconsistent, because at least one of them must be false. Once you think you're wrong, you must be right! Crucial to my argument is a distinction between explicitly contradictory beliefs and three forms of inconsistency in belief. In the first of these, you believe contingent propositions each of which is logically independent of the others, yet you also believe the syntactic negation of their conjunction. In the second, your beliefs are inconsistent because of the rigid designation of demonstratives such as 'this' embedded in their contents. In the third, as in Modesty and Modesty* your beliefs are inconsistent because of self-reference.

Keywords Inconsistency · Contradiction · Rationality · Self-reference

✉ John N. Williams
williamsjohn276@gmail.com

¹ School of Creative Arts and Humanities, Charles Darwin University, Darwin, Australia

In D.C. Makinson's original version of the preface paradox, you write a long book consisting of statements $S_1, S_2 \dots S_n$. Let us suppose, along with Makinson, that each of your beliefs in each statement in the corpus of your book is justified.¹ Although he does not say as much, these might be contingently true or false. But now you are struck by the thought that you have written many similar books before, each time only to later discover errors. This seems evidence enough to claim in the preface that there is indeed an error in the corpus of your book. So you appear to be justified in believing that $\sim(S_1 \& S_2 \& \dots S_n)$. As Makinson puts it

... to say that not everything I assert in this book is true, is to say that at least one statement in this book is false. That is to say that at least one of s_1, \dots, s_n is false, where s_1, \dots, s_n are the statements in the book; that $\dots \sim(s_1, \& \dots \& s_n)$, is true. (1965, 205).

But now you appear to have inconsistent beliefs because the conjunction of their contents i.e. $[S_1 \& S_2 \& \dots S_n \& \sim(S_1 \& S_2 \& \dots S_n)]$ is necessarily false. Yet each of your beliefs is justified. Makinson concludes that this is 'a living and everyday example' in which rationality allows you to hold 'logically incompatible beliefs' even if you *know* that that your beliefs are so (1965, p. 205, my italics), in contradiction of orthodox views of justification (Chisholm 1989; Hempel 1962; Lehrer 1974; Pollock 1983; Quine and Ullian 1970; Schick 1963; Swain 1970).

Perhaps the most vigorous opposition to his conclusion comes from classical and insightful objections by Doris Olin (2003). Here I argue for the more radical conclusion that there are 'living and everyday' examples in which rationality may *require* you to have inconsistent beliefs, even when you recognize the inconsistency. I defend this argument against objections, including Olin's.

Here is how I will proceed. In the next section I set out my assumptions and distinguish 'explicitly' contradictory beliefs (those that have contents that are syntactic negations) from inconsistent ones. I also demarcate two forms of inconsistency in belief that will play no role in what follows. In Sect. 2, I distinguish three types of inconsistency in belief that I will go on to exemplify with several versions of the paradox. I argue that these types of inconsistency in belief are epistemically better than explicitly contradictory beliefs. In Sect. 3 I return to Makinson's original version of the paradox. I consider several objections to it (Ryan 1991; Lacey 1970; New 1978; Olin 2003) concluding that each fail. However, there is a problem with it that seems to have gone unnoticed, namely that it is unrealistic to expect that one could have the beliefs needed for the paradox to get off the ground. There is a way of reformulating it to avoid this problem, but then it is doubtful that the case counts as a living and everyday example. In Sect. 5, I give *World Capitals*, a version that evades this difficulty. However, I show in Sect. 6 that this version is vulnerable to one of Olin's objections, one she makes against her own version of the paradox, *Fallibility*. In Sect. 7, I present *Modesty*, a version that supposes that you believe that at least one of your beliefs (*excluding* this) is false. I argue that this version escapes all the objections that could trouble the other versions as well as some interesting general objections that Olin makes. I conclude that this is a living and everyday case in which

¹ Makinson appears to assume this. He nowhere defends it.

rationality requires you to have inconsistent beliefs even while you recognize that your beliefs are inconsistent. In Sect. 8, I argue more tentatively for the same verdict for *Modesty**, a version that supposes that you believe that at least one of your beliefs (*including this*) is false.

1 Preliminary remarks

Here are some preliminary clarifications. I take a paradox to be a set of premises that seem obviously true that lead with apparently impeccable logic to a conclusion that seems obviously false.² Of course, one solution to a paradox is to show that despite first appearances and perhaps even counterintuitively, the conclusion is true. If this succeeds, then we will no longer think of the argument as a paradox.

I take rationality in belief to be a matter of following correct rules for forming beliefs. For instance, rationality will instruct us to gather evidence, to form beliefs if they are properly justified by it, to not believe what we can be expected to see must be false and to be prepared to revise our beliefs. Throughout I will talk of justification as evidence.

My conclusion is not meant to be about a perfectly logical thinker who collects and uses evidence perfectly and whose cognitive faculties are all unlimited. It is meant to be about real human thinkers such as you and me.³ Thus arriving at this conclusion might give us interesting insights into ourselves. That said, we can aim for consistency among our beliefs. Were we to attain near-perfect success in this, then no doubt we would be exceptional, but would still be recognizably human.

One important human limitation that my argument will rely upon is a limitation of human thought. This limitation is exemplified by the fact that you may be able to think thoughts that I cannot think, or vice versa. This becomes starker if we suppose that we are at different stages of development or belong to different cultures or even different species. One explanation of the difference is that one of us has concepts unavailable to the other. Another is that one of us can think more complex thoughts than the other, even when the same concepts are available to us. In the light of this limitation, *Searle's principle* is plausible, as follows.

If one has a belief, then one has the ability to think the thought of its content. (Searle 1992, pp. 155–162).⁴

I will be partly concerned with what are often called ‘occurrent’ beliefs. In forming such a belief, you think the thought of its content and mentally assent to that thought. For example, you might see no traffic on a road that you wish to cross. In thinking

² In this I follow Sainsbury (1995, p. 1). Olin (2003, pp. 6–7) calls this a ‘type 1 paradox’, a type 2 paradox being a set of premises that seem obviously true that lead with apparently impeccable logic to multiple conclusions that seem obviously inconsistent. All five versions of the paradox that I will discuss are of type 1.

³ In this respect I follow those such as Foley (1992), rather than those such as Ellis (1979), Hintikka (1962) and Stalnaker (1987).

⁴ For objections to Searle’s principle, see Burge (1978, p. 133) and Sorensen (1985, p. 496). For replies, see Williams (2007, p. 150, note 15).

the thought that it is safe to cross and in mentally assenting to it, you believe that it is safe to cross. Then there are what we might call ‘dispositional beliefs’. For example, before reading this sentence you probably knew, and so believed, that the Earth is the third planet from the Sun although you did not believe this occurrently.⁵

Searle’s principle helps to explain an important property of belief. This is that belief distributes but does not always collect over conjunction. More formally, what seems correct is *belief-distribution*, as follows.

If you believe that (p and q), then you believe that p . And you believe that q .⁶

To illustrate this, suppose that you believe that it is raining in London and it is cold in London. Then you believe that it is raining in London. And you believe that it is cold in London. Belief-distribution appears to be immune from counterexample.⁷ What doesn’t seem correct is *belief-collection*, as follows.

If you believe that p and you believe that q , then you believe that (p and q).

The principle is dubious as applied to beliefs that are unconscious, namely those you have but are unaware of having. Many of your perceptual beliefs are unconscious and are in constant flux, coming and going in step with changes in how things seem to you. Surely you cannot be aware of holding a belief that conjoins the contents of beliefs of which you are unaware. For example, in watching a sunset, you are normally unaware of a rapidly changing conjunctive belief. Nor does it seem likely that you always have that belief unconsciously.

Belief-collection also seems implausible for beliefs in general. If the set of your occurrent and dispositional beliefs is large enough, it is plausible that you might be unable to think the thought of the conjunction of their contents, not because you lack the concepts needed to think the would-be thought, but because that thought is just too complex for you to think. In that case you could not believe the conjunction. You would be a bit like a child who having forgotten to bring her mother’s shopping list, can remember to buy each item on seeing it but is unable to remember the list.

Since you might have a set of beliefs without being *able* to believe the conjunction of their contents, it follows that even if you have justification for believing that conjunction, you are unable to form a justified belief in it. Thus despite the justification you have, you have no epistemic obligation to believe it, at least not if ‘ought’ implies ‘can’. Surely rationality cannot require you to believe what is humanly impossible for you to believe. It will turn out that this will present an obstacle to some versions of the preface paradox, depending on the size of the set of your beliefs.

A distinction between *explicitly contradictory* beliefs and *weakly inconsistent* beliefs will be crucial to my argument. You have contradictory beliefs just in case

⁵ Following Audi (1994), dispositional beliefs should be distinguished from dispositions to believe. For example, while asleep last night you neither believed, nor therefore knew, that there were no flying elephants in view. Yet you were disposed to have that belief upon considering the proposition that there are no flying elephants in view. I am grateful to a first anonymous referee for pointing this out.

⁶ To account for temporality, we could formulate this as follows. If at t you believe that (p and q), then at t you believe that p . And at t you believe that q .

⁷ A lone attempt to counterexample belief-distribution is made by Pruss (2011). For a reply, see Williams (2014, p. 4, note 3).

it is logically necessary that the content of one is true and that of the other is false. This is contradiction in terms of *classical* logic. I mean to exclude a case in which the content of one is true *and* false. It follows trivially that you have a *pair* of beliefs. The simplest case would be your belief that p paired with your belief that not- p , with the two contents being syntactic negations of each other. An example is your belief that all logic books are boring, paired with your belief that not all logic books are boring. Let us call this ‘explicit’ contradiction in belief. In contrast, your belief that only boring books are logic books contradicts your belief that not all logic books are boring, but not explicitly, because their contents are not syntactic negations. We might call this ‘implicit’ contradiction in belief. In considering the contents of a pair of explicitly contradictory beliefs, the only feature of these that needs your attention in order to recognize the contradiction is the syntactic negation. You should recognize that a belief that all logic books are boring and a belief that not all logic books are boring cannot both be true, although one must be true.

In contrast, it might be more difficult for you to recognize that ‘only boring books are logic books’ contradicts ‘not all logic books are boring’. Students beginning to study logic commonly find this difficult to see. It is plausible that to the extent that this is more difficult for you to see, the more you are to be excused in holding such a pair of beliefs. Suppose for example that having just measured yourself, you believe (while conceiving of yourself in first-personal terms) that you are not taller than 6 feet. But on the basis of reliable testimony, you also believe that your mother’s brother’s only nephew is taller than 6 feet. You might be forgiven for not realising that this relative can only be you.

It is worth noting that if more beliefs are added to a set of explicitly contradictory beliefs, then that set ceases to be a set of explicitly contradictory beliefs, since it remains such a set only so long as it constitutes a pair of beliefs. It would however, have a set of explicitly contradictory beliefs as a proper subset.

Let us also say that your beliefs are inconsistent just in case it is logically impossible for all their contents to be true (Olin 2003, p. 61; Foley 1979, p. 247). Thus if you have explicitly contradictory beliefs then you have inconsistent ones. The converse need not hold however.

To see this, consider a pair of beliefs that are not in contradiction but in which the content of one entails the negation of the content of the other. This might be a case in which you believe that p and you believe that q , but although these two contents are not contradictory, still if p then not- q . For example, suppose that you believe that your meal contains tomatoes. You also believe that it contains no fruit. But if it contains tomatoes then it contains some fruit, because tomatoes are fruit. Your beliefs are not in contradiction, because your meal might contain only bananas, in which case both your beliefs are false. But you might not know that tomatoes are fruit, thinking mistakenly that they are vegetables. Surely you need not believe that your meal contains tomatoes and also believe that it does not. Nor need you believe that it contains some fruit and also believe that it does not.

Having made this point, I should now say that my argument will not concern itself with this particular form of inconsistent beliefs. One reason is that you may be excused in not recognising the inconsistency. More importantly, if you do recognize

it, then surely rationality would instruct you to revise your beliefs. Once you learn that tomatoes are fruit, you should give up at least one belief.

There is another form of inconsistent beliefs as well. This is a set of beliefs that includes at least one belief the content of which is a necessary falsehood. Mathematicians before Gödel were not irrational in believing that arithmetic is decidable, because they could not have been expected to know that it is necessarily false that arithmetic is decidable (de Almeida 2001, pp. 39–43). My argument will not concern itself with this particular form of inconsistent beliefs either. One reason is that it is not a set of contingent beliefs and I will follow Makinson in considering versions of the paradox that only involve contingent beliefs. Another is that it is not a case in which one has a set of inconsistent beliefs while *recognizing* the inconsistency.

2 Three other types of inconsistency in belief: and why these are epistemically better than explicitly contradictory beliefs

I will be concerned with three types of inconsistency in belief. The first type is found in a case in which you believe each of a number of contingent propositions each of which is logically independent of the others, yet you also believe the syntactic negation of their conjunction. We may represent this as a case in which you believe that p_1 , you believe that p_2 ... you believe that p_n , and you believe that $\text{not}(p_1 \ \& \ p_2 \ \& \ \dots \ p_n)$. The inconsistency arises purely from the syntax of the conjunction of the contents of your beliefs. For want of a better term, let us call these ‘weakly inconsistent’ beliefs.

In contrast, there are two other types of inconsistency in belief in which the source of inconsistency is not purely syntactic. As a case of the first type, suppose that you are searching for a small ball suitable to amuse your baby. You empty a jam jar of 20 assorted small balls into a tub. You look at the balls and think ‘One of *these* is suitable’. Then you inspect each ball in turn, each time thinking ‘*That* is unsuitable’. Both ‘*That* is unsuitable’ and ‘One of *these* is suitable’ rigidly designate *actual* balls in the tub via the demonstratives ‘that’ and ‘these’. You are not thinking of balls that *might* be in the tub. For ease of exposition let us number the contents of your beliefs in the unsuitability of each ball, calling the first time you think ‘*that* is unsuitable’, ‘*that*₁ is unsuitable’ and the second time you think it, ‘*that*₂ is unsuitable’ and so on. Your beliefs are inconsistent because it is logically impossible that (*that*₁ is unsuitable & *that*₂ is unsuitable & ... *that*₂₀ is unsuitable & one of *these* is suitable), given the rigid designation involved. For want of a better term, let us call this ‘demonstrative’ inconsistency in belief.

As a case of the second type, suppose that I believe that my total beliefs—excluding this one—contain at least one false belief. Suppose also, as is entirely realistic, that I have ‘other’ beliefs as well. My second-order belief is self-referential, owing to the demonstrative ‘this one’. However, its content excludes itself as a candidate for falsehood. For want of a better term, let us call this ‘exclusively self-referential’ incon-

sistency in belief.⁸ My total set of beliefs is inconsistent, because if my second-order belief is true then at least one of my other beliefs must be false.

Let us first compare explicitly contradictory beliefs with each of these three types of inconsistent belief in terms of the ratio of your true beliefs to your false ones. In believing that p and also believing that not- p , you are guaranteed to have one true belief and one false belief. The same is true of the weakly inconsistent set. If your belief that not- $(p_1 \& p_2 \& \dots p_n)$ is true, then at least one of your beliefs that p_1 , that p_2 , ... that p_n must be false. You are also guaranteed to have at least one true belief. For if all of your beliefs are false, then not- $p_1 \& \text{not-}p_2 \& \dots \text{not-}p_n \& (p_1 \& p_2 \& \dots p_n)$, which is impossible.

The same is true of your beliefs about the balls. As we have just seen, you must have at least one false belief. You must also have at least true one, because if all your beliefs are false then *that*₁ is suitable & *that*₂ is suitable & ... *that*₂₀ is suitable & one of *these* is unsuitable. This is impossible given the rigid designation involved. Likewise, suppose while conceiving of yourself in the first-person, you believe that your total beliefs—excluding this one—contain at least one false belief. You are guaranteed to have at least one true belief because if all your other beliefs are false, then it is true that your total beliefs—excluding this one—contain at least one false belief.

But unlike the case in which you have explicitly contradictory beliefs, in the other three cases most of your beliefs could be true. Take your weakly inconsistent beliefs. It might turn out that not- $p_1 \& p_2 \& \dots p_n \& \text{not-}(p_1 \& p_2 \& \dots p_n)$. As n becomes larger, so the set of these possibly true beliefs increases. Likewise, your only false belief about the balls might be that *that*₁ is suitable. Likewise again, only one of your other beliefs—excluding this—might be false. You should take this scenario seriously, because you might have no idea which of your beliefs is false. You have a much better chance of maximizing the ratio of your true beliefs to your false ones than you do in having explicitly contradictory beliefs, especially if many of your beliefs are involved.

Now let us compare the cases in terms of evidence. First consider the case in which you believe that p and you believe that not- p . Any evidence that you have for the content of either belief is evidence against that of the other. You cannot be justified in both beliefs. Contrast this with the case in which you believe that p_1 , you believe that p_2 ... you believe that p_n , and you believe that not- $(p_1 \& p_2 \& \dots p_n)$. Good evidence against the conjunction $(p_1 \& p_2 \& \dots p_n)$ need not be good evidence against any particular conjunct. If your evidence against any conjunction must also be considered evidence against each conjunct, then in some cases, to the degree it counts against one conjunct, it will count equally against every conjunct.

Now suppose in such a case that you do have good evidence against the conjunction. Thus if your evidence is good enough to compel you to give up your belief in it, then it is good enough to compel you to give up your belief in every other conjunct as well. But surely you need not do that. After all, you have no idea which conjuncts are false and as we have just noted, for all you know, most and many of them are true.

The same argument applies to the case of the balls, provided that any particular ball's unsuitability does not entail the suitability of any other. This proviso could be

⁸ I might instead believe that my total beliefs—including this one—contain at least one false belief. We could say that this involves 'inclusive' self-reference. In Sect. 7 I will argue that this too involves inconsistency.

met realistically. A ball could fail to amuse your baby in many unrelated ways, in terms of properties such as colour, size, texture, material, elasticity or plasticity. This argument extends to the case in which you believe that your total beliefs—excluding this one—contain at least one false belief, provided that your other beliefs are consistent with each other. Even if they are not entirely consistent, it is still plausible that your evidence for your second-order belief need not count against any these other beliefs. This is because that evidence might be just induction from your knowledge that your beliefs have always turned out in the past to contain falsehood. In contrast, evidence for each of your other beliefs will realistically include evidence of *every kind*. Likewise, the evidence for each of $p_1 \& p_2 \& \dots p_n$ might be of *different kinds* from the evidence against the conjunction, a point to which I will return in the next section. Likewise again, your evidence for your belief that one of the balls in the tub is suitable might be that you seem to remember putting a ball that amused your baby on many occasions into the jam jar, although you now have no idea of what it was like. In contrast, your evidence that particular balls in the tub are unsuitable might be deduction or induction from your perception of their properties plus what you know of your baby's preferences.

So it appears that you might have good evidence against $(p_1 \& p_2 \& \dots p_n)$ and yet still have good evidence for each conjunct. Likewise, it appears that you could have good evidence that one of the balls in the tub is suitable and yet have good evidence that each particular one is unsuitable, or have good evidence that your total beliefs—excluding this one—contain at least one false belief and yet have good evidence for each of your other beliefs. In terms of evidence, you are much better off than having explicitly contradictory beliefs.

It even seems that in each of these three cases of inconsistent belief, all but one of your beliefs could constitute knowledge. All but one might be true, justified and not Gettiered. Or to avoid taking sides on externalism about knowledge, why couldn't all but one of your beliefs track the truth or be safe true beliefs? In contrast, when you have explicitly contradictory beliefs, you may know at most only half of what you believe.⁹

Finally, let us compare the cases in terms of how you should revise your beliefs once you recognize the inconsistency that they involve. Suppose that you come to recognize that you both believe that p and also believe that not- p . You need no evidence at all to see that you must have a false belief about whether or not p . You should recognize this. You know exactly what you must decide, namely whether or not p . Rationality tells you to look again at the evidence you have for and against p and if necessary, gather more. If that justifies you in keeping one of the two beliefs then you must give up the other. If this does not help, then you must suspend judgement and give up both beliefs. Either way, you must revise your beliefs. If you do nothing, then rationality will find you at fault.

⁹ Given *anti-incoherence*, namely that one cannot believe the negation of what one knows, it follows that one cannot know anything of what one believes. Claudio de Almeida endorses a related principle: if one believes that p and also believes that not- p , then one knows neither, nor is one justified in believing either (de Almeida 2012, p. 205). I hesitate to accept anti-incoherence. Why couldn't you be self-deceived in believing what you know to be false?

In contrast, suppose again that you believe that p_1 , you believe that p_2 ... you believe that p_n , and you believe that $\text{not}-(p_1 \ \& \ p_2 \ \& \ \dots \ p_n)$. Suppose also that you come to recognize that at least one of your beliefs must be false. What action should you undertake? The way forward is much murkier than it was in the case in which you recognize that your beliefs are in explicit contradiction. Your problem is that recognizing the inconsistency gives you no clue as to which belief is false. For all you know, they might all be true except one. On the other hand, if merely recognizing the inconsistency were to give you reason to suspend judgement in any particular one of p_1, p_2, \dots, p_n , then it would give you equally good reason to suspend judgement in any other. But that would be hasty, because suspending judgement in each might appreciably lower the ratio of your true beliefs to your false ones, and you might have good independent evidence for each. You do have the incentive to re-evaluate the evidence for and against each. But if that does not significantly change the balance of evidence, then it seems that rationality provides you with no guidance at all about how to revise your beliefs. But if so, then it cannot find you at fault for keeping all your beliefs.

Much the same can be said for demonstrative inconsistency. Suppose that you realize that it must be false that each ball in the tub is unsuitable and also that one of these is. That all by itself gives you no clue about which of your beliefs are false. If it were to give you reason to suspend judgement in any ball's suitability then it would give you equally good reason to suspend judgement in the suitability of any other. Likewise, suppose that you realize that your belief that your total beliefs—excluding this one—contain at least one false belief, means that your total beliefs are inconsistent. If this realization were enough to compel you to suspend judgement on any particular matter in which you now believe, then it would be enough to compel you to adopt a total suspension of belief. For even if some of your other beliefs are inconsistent, still your realization of inconsistency in your total beliefs all by itself gives you no reason to single out any particular one of your other beliefs as being false rather than any other. In terms of rationality, you are much better off than in the case in which you have explicitly contradictory beliefs.

In what follows I will discuss versions of the paradox that exemplify these three forms of inconsistent beliefs. I am now finally in a position to discuss the various versions of the preface paradox, first returning to Makinson's.

3 Makinson's original version of the paradox

As we saw in the introduction, Makinson reaches the conclusion that

... to say that not everything I assert in this book is true, is to say that at least one statement in this book is false. That is to say that at least one of s_1, \dots, s_n is false, where s_1, \dots, s_n are the statements in the book; that $\dots \sim (s_1, \ \& \ \dots \ \& \ s_n)$ is true. (1965, p. 205).

But now you appear to have weakly inconsistent beliefs because the conjunction of your beliefs, i.e. $[S_1 \& S_2 \& \dots S_n \& \sim(S_1 \& S_2 \& \dots S_n)]$ is necessarily false. Yet each of your beliefs is justified.

There are only three strategies for avoiding this result. One may deny that your beliefs are inconsistent, one may deny that that are all justified or one may deny that you are able to form the beliefs in question.

Following the first strategy, A. R. Lacey claims that “modesty ... demands no more in a preface than the statement ‘*Probably* there are errors in my book” (1970, p. 614). This makes your beliefs consistent, because it is logically possible that $[S_1 \& S_2 \& \dots S_n \& \text{probably } \sim(S_1 \& S_2 \& \dots S_n)]$. This however, only changes the original example. Surely it is possible for an author to claim sincerely that there are, or even that there certainly are, errors in her book. For example, John Passmore says in his preface to *A Hundred Years of Philosophy* ‘This book contains a large number of errors ... simple slips and plain mistakes. So much I know a priori, but not, of course, what they are’ (1957, p. 8). One cannot deny the possibility inherent in the example merely by changing it.

Sharon Ryan (1991) claims that given the evidence $E_1, E_2 \dots E_n$ for each of the statements in your book, the most that you are justified in believing is that your book *might* contain an error, not that it *does* contain one. This again makes your beliefs consistent, because it is logically possible that $[S_1 \& S_2 \& \dots S_n \& \text{possibly } \sim(S_1 \& S_2 \& \dots S_n)]$.

Against this, consider again the possibility that you have written a very large number of similar books in the past, using similar methods, only to later discover that it contains at least one error. We are of course free to stipulate this number to be as large as we please. This appears to provide us with inductive justification that is as strong as we please for thinking that your present book also contains at least one error.

To insist that this inductive basis of yours supports only the different conclusion that your book probably contains error is to quarrel with the classical view of the very nature of induction. After all, induction is classically thought of as adducing evidence that licenses the conclusion that a proposition is true, not just probably true, although of course the evidence does not provide a deductive guarantee of the conclusion. We do not think of induction as deduction from the evidence to the conclusion that a proposition is probably true. Nor do we think of it as adducing evidence that non-deductively licenses the conclusion that a proposition is probably true. So given this classical perspective, you should believe that your book *does* contain at least one error.

Saying that is consistent with a different verdict in a case in which you know only that your ticket is one of very many, say one million, in fair lottery. In other words, you know that there is a very high statistical probability that your ticket will lose. Some may think that you should believe only that, and refrain from believing that it *will* lose.¹⁰ The two cases are different, because the lottery case makes no appeal to induction. But there do seem to be cases in which your knowledge of very high statistical probability alone suffices to justify a belief that a proposition is true. As an analogous case, suppose that I acquire the irrational fear that all the molecules of oxygen in the room will suddenly congregate in one corner, resulting in our suffocation.

¹⁰ Against Klein (1985, pp. 120–121).

Knowing only that there is roughly one in a googplex chance that this will occur, you tell me solemnly, ‘That won’t happen’. Your sincerity is not only possible, but also proper. Indeed, it seems improper of you to tell me instead ‘That’s very unlikely to happen’ because that would give me less reassurance than is appropriate. You stand by truth, not wishy-washy probable truth.

Another source of justification lies in your knowledge that given human fallibility, the longer the book, the greater the chance of error. For a real case, Hazel Barnes says in her translation of Sartre’s *Being and Nothingness* ‘In a work *as long as this* there are certain to be mistakes’ (1958, p. vii, my italics).

Another probabilistic way of denying that your beliefs are inconsistent is to insist that the evidence $E_1, E_2 \dots E_n$ for each of the statements in your book justifies you only in believing that each statement is probably true. Once again your beliefs are consistent, this time because it is logically possible that [probably S_1 & probably S_2 & ... probably S_n & $\sim(S_1 \& S_2 \& \dots S_n)$]. For example, Christopher New claims that evidence that your present book contains error is relevant to each of your beliefs $S_1, S_2, \dots S_n$ because it reduces the probability of each of $S_1, S_2, \dots S_n$ equally (1978, p. 343).

But we just saw in the last section that good evidence against a conjunction need not be good evidence against any of the conjuncts. So even if this evidence does reduce the probability of each of $S_1, S_2, \dots S_n$, it does not compel you to give up any of your beliefs in these individual statements. Perhaps the reduction would be miniscule. Secondly, a conjunction can be improbable relative to a set of evidence without its conjuncts being improbable relative to that same evidence. If my evidence for the proposition that Carlos and Natasha did not both murder Senator Linda is that in numerous past assassinations both have always preferred to work alone, then relative to that information it is improbable that Carlos and Natasha both murdered her. But this information does not decrease the probability that Carlos murdered her nor that Natasha did. As another case, if three balls are taken at random each from three separate urns, each containing 50 white and 150 black balls, then relative to the information that this has been done, the probability that all the three balls are black is low (i.e. $3/4 \times 3/4 \times 3/4$), i.e. $27/64 = 0.42$). But relative to the same information, the probability of each ball not being black is still high, i.e. $3/4 = 0.75$.

Besides, the evidence you have against the conjunction of statements in your book is not relevant to the truth of any of these statements because that evidence is of a radically different kind from that which justifies each statement. Suppose that your book contains only historical claims. Evidence for each statement will include records, testimony and archaeological results. In contrast, your evidence against their conjunction might be purely inductive, comprising the fact that you have written very many similar historical books that you later discovered to contain error. Or suppose that you have evidence that *ceteris paribus*, the longer a list of statements, the greater the likelihood of error. This is the kind of evidence which any proof-reader or computer programmer will adduce from experience. The evidence might even take the form of a comparison of the number of errors in short lists with that in long ones. Call this evidence simply ‘E’. Once again this is an entirely different kind of evidence from any that might undermine particular historical claims made in your present book.

Such insistence on the dependence of the available sets of evidence may stem from a crucial failure to recognise that the relevant beliefs are *weakly inconsistent* rather than *explicitly contradictory*. New’s view seems plausible only so long as one assimilates the justification of believing *each* of the statements with the justification of believing *all* of them; the source of the paradox, such as it is, is *not* that we are forced to take explicitly *contradictory* attitudes towards the *conjunction* of statements in the book (as supposed by Hoffman 1968, p. 122; Routley and Routley 1975, p. 211), i.e. *not* that ‘in the light of $[E_1, E_2 \dots E_n]$ we would be right to believe the conjunction [of statements] but in the light of $[E]$ we would be right to disbelieve it’ (New 1978, p. 343).

The original case is not one in which you believe the conjunction of statements in the corpus of your book, therefore you do not hold explicitly contradictory beliefs, but weakly inconsistent ones. If you believe that S_1 and believe that S_2 and ... and believe that S_n and yet believe that $\sim(S_1 \ \& \ S_2 \ \& \ \dots \ \& \ S_n)$, then you do not hold explicitly contradictory beliefs (as you would if you believed that $(S_1 \ \& \ S_2 \ \& \ \dots \ \& \ S_n)$ and also believed that $\sim(S_1 \ \& \ S_2 \ \& \ \dots \ \& \ S_n)$), since there are no two beliefs that contradict each other. Yet your beliefs are inconsistent because it is logically impossible that they are all true.

The crucial distinction between weakly inconsistent beliefs and explicitly contradictory ones is importantly relevant to the question of independence of evidence and suggests that the apparently paradoxical conclusion—that you have inconsistent beliefs each of which is rational—appears unacceptable only because it is liable to be mistaken for a different conclusion that you have explicitly contradictory beliefs each of which is justified. As we saw in the last section, that conclusion is indeed unacceptable and yet rationally does not appear to prohibit you from having weakly inconsistent beliefs. Indeed, it seems possible that you may have good justification for each belief and that your justification for each is not justification for the falsehood of any other.¹¹

Olin gives a different kind of objection to the argument of the original preface paradox, namely that it proves too much. She writes as follows.

¹¹ The failure to recognize this distinction is found in a recent approach by Hannes Leitgeb. He considers scientists who publish experimental results $A_1 \ \& \ A_2, \dots, A_n$. Leitgeb proposes that they have not asserted the conjunction of these statements. Although ‘if one utters a *single* declarative sentence A, \dots one asserts that A (2014, p. 13, my italics), in the context in which there are a ‘great many’ of these statements made in ‘one fell swoop’, they have only asserted that ‘the vast majority of the data are correct’, where “the exact meaning of ‘vast majority’ is determined partially by the context” (2014, p. 12).

Whatever one may think of this proposal, accepting it does not remove the paradox. Leitgeb thinks of the paradox as a case in which rationality requires that you have *explicitly contradictory beliefs*. His solution goes as follows. In justifiably asserting each of A_1, A_2, \dots, A_n in your book, you justifiably assert their conjunction, i.e. $A_1, \ \& \ A_2 \ \& \ \dots \ \& \ A_n$. But in your preface, you justifiably assert that $\sim(A_1, \ \& \ A_2 \ \& \ \dots \ \& \ A_n)$. If all your assertions are sincere then you have explicitly contradictory beliefs, each of which is justified. Leitgeb proposes that in asserting each statement in the book you do not assert their conjunction but only assert that the vast majority of these are true. So ‘No contradiction arises’ (2014, p. 14).

But Leitgeb does not deny that you assert *each* of A_1, A_2, \dots, A_n in your book, so you have still made a set of assertions $A_1, \ \& \ A_2 \ \& \ \dots \ \& \ A_n \ \& \ \sim(A_1, \ \& \ A_2 \ \& \ \dots \ \& \ A_n)$. Although none of these *contradicts* any other, your assertions are *inconsistent* because it is impossible that all of them are true, something you should recognize. If each is sincere then you have inconsistent beliefs, but since each belief is justified by independent evidence, each is rational. Therefore Leitgeb’s proposal leaves the paradox untouched.

Readers, as well as authors, can employ the inductive inference based on past errors. If I know that the statements in David's book are D_1, \dots, D_n , then I can argue, based on the past fallibility of David and other authors, that $\sim(D_1 \& \dots \& D_n)$; if I know that the statements in Jennifer's book are J_1, \dots, J_n , then I can justifiably infer that $\sim(J_1 \& \dots \& J_n)$; and so on. This is not the sort of information one would normally expect to acquire from reading careful and serious books (2003, p. 65).

But this way of describing matters makes the conclusion of the argument seem more counterintuitive than it really is. It is not one's *reading* of the books, whether or not they are careful and serious, that provides one with justification for thinking that each contains at least one error. One does not need to read them. All one needs is to know is that David has previously written *very many similar* books, using *similar methods*, which contain at least one error, and that his present book contains a *very large* number of statements. That is enough to justify one in believing that his present book contains at least one error. The same is true of Jennifer, or any similar case.

However, there is an important mistake in Makinson's logic that seems to have long gone unnoticed. This is that your belief that there is indeed an error in the corpus of your book is in fact *consistent* with your beliefs that S_1 , that S_2, \dots that S_n , because it is *possible* that there is a false statement in your book *other than* $S_1 \& S_2 \& \dots S_n$, each of which is true. In other words, the conjunction of the contents of your beliefs, i.e. $(S_1 \& S_2 \& \dots S_n \& \text{there is an error in the book})$, is possibly true.¹²

One way to regain inconsistency is to stipulate that you indeed sincerely assert in your preface, 'It is false that $(S_1 \& S_2 \& \dots S_n)$ '. This however has the strange and very unrealistic consequence that now your preface might be longer than your book! To appreciate this point, let us denote every statement in Passmore's *A Hundred Years of Philosophy* ' S_1 ', ' S_2 ', and so on. Given unrealistically that *A Hundred Years of Philosophy* minus its preface consists entirely of statements, its preface is now *A Hundred Years of Philosophy* again but prefixed by 'It is false that'. Even if the vast majority of the corpus of the book is comprised of statements rather than questions and suchlike, the preface will still be a hefty beast. So we have strayed from the *desideratum* that the case should be a 'living and everyday example'.

A more important difficulty is that we might doubt that it is humanly possible for one to believe something of the form ' $\sim(S_1 \& S_2 \& \dots S_n)$ ' for a book of any appreciable length. Could Passmore really form a belief about the conjunction of *all* the statements in *A Hundred Years of Philosophy*? It seems unlikely that *any* of us could think the thought of the conjunction, in which case Searle's principle prohibits any one of us from believing it.

This suggests recasting the original paradox in terms of demonstrative reference. Suppose that you claim sincerely in the preface that you believe each statement in your book and that at least one of these is false. In believing that at least one of *these* is false you need not think of the negation of a conjunction of statements. Nor need you identify each statement in your book as belonging to those in your book or even know whether each is one of these. Just as 'One of *these* is suitable' rigidly designates the

¹² So against Makinson, where s_1, \dots, s_n are the statements in the book, to say that at least one statement in this book is false is *not* to say that $\dots \sim(s_1, \& \dots \& s_n)$ is true.

set of actual balls in the tub, so ‘at least one of *these* is false’ rigidly designates the set of actual statements in your book. You are not thinking of statements that *might* be in your book. Now your beliefs are inconsistent, because it is logically impossible that ($S_1 \& S_2 \& \dots S_n \& \text{At least of one } these \text{ is false}$), where ‘these’ rigidly designates $S_1 \& S_2 \& \dots S_n$.¹³

Yet a worry remains. For the case to succeed, each of your beliefs in S_1 and in $S_2 \dots$ and in S_n must be justified. The obvious strategy for meeting this requirement is to suppose that they are consistent, otherwise evidence for some will be evidence against others. On the one hand, an author of a consistent book would be recognizably human. On the other hand, we might doubt that a book consisting entirely of a set of consistent statements rather than questions and suchlike counts as a ‘living and everyday example’ of a book. A list, rather than a book, seems a better candidate for such a set. This suggests a second version of the paradox.

4 A second version of the preface paradox: world capitals

Consider *World Capitals*. Tonight is quiz night at your local pub and the winner of this week’s quiz is the person who can correctly name the most capitals of the countries of the world. To prepare for the quiz you have tried to memorize the capital cities of all 196 countries in the world by repeatedly writing a list ‘The capital of Afghanistan is Kabul, the capital of Albania is Tirana, the capital of Algeria is Algiers, ... the capital of Zimbabwe is Harare’. You believe each of these 196 contingently true statements on the conjunctive evidence that you seem to remember that it is contained in a second list that you know was written by an expert geographer and also that you know that your memory is very accurate. This memorial-cum-testimonial evidence is the *same* for each of your 196 beliefs. The truth (or falsehood) of any one of them does not entail the falsehood (or truth) of any other. So your justification for the truth of each belief is not justification for the falsehood of any other.

Yet you also have justification for believing that there is an error in your present list, namely that on many occasions on which you have written the list and compared it to the second, you have discovered that you have misremembered at least one capital city, although you now have no idea which these were. This is inductive, not memorial-cum-testimonial evidence. We have already established that this does not count against any of the 196 statements, nor undermine the memorial-cum-testimonial evidence you have for each. When you are called upon in the pub to name the world capitals, you sincerely assert each of the 196 statements ‘The capital of Afghanistan is Kabul, the capital of Albania is Tirana, the capital of Algeria is Algiers, ... the capital of Zimbabwe is Harare’. Then you ruefully but sincerely add ‘But I’m sure that at least one of these statements is wrong’.¹⁴

The demonstrative element ‘these statements’ of the content of your occurrent belief ensures inconsistency, because it is logically impossible that (The capital of

¹³ I thank a second anonymous reviewer for persuading me of this point.

¹⁴ Alternatively, to make the inconsistency more salient, you might add instead ‘At least one of these statements is wrong. I’m sure of that’ or just ‘At least one of these statements is wrong’.

Afghanistan is Kabul, the capital of Albania is Tirana, the capital of Algeria is Algiers, ... the capital of Zimbabwe is Harare & at least one of *these* statements is wrong). While it seems counterintuitive that you could believe the conjunction of all the statements in *A Hundred Years of Philosophy*, it seems not at all counterintuitive in *World Capitals* that you could believe that at least one of *these* statements is wrong. You could have the ability to think thoughts of *these* statements, namely *those you have just made*, even if you did not have the ability to think the thought of their conjunction. 'At least one of *these* statements is wrong' rigidly designates the set of statements that you have actually just made. You are not thinking of statements that you *might* have just made.

Alternatively, we could ensure inconsistency by supposing that your sincere and rueful addition is instead 'But I'm sure that at least one of *these 196 statements* is wrong'. Part of your justification comes from knowing that there are no more than 196 statements that you have tried to memorize. Since in fact you have made exactly 196 assertions about capital cities, it is logically impossible that (The capital of Afghanistan is Kabul, the capital of Albania is Tirana, the capital of Algeria is Algiers, ... the capital of Zimbabwe is Harare & at least one of *these 196 statements* is wrong). In fact, that just-bracketed content contains exactly 197 statements.

But while your rueful addition might occasion you some embarrassment, it does not seem to constitute any irrationality on your part. Knowing of each of the 196 statements that you seem to remember that it is contained in a second list that you know was written by an expert geographer and that you know that your memory is very accurate, is excellent justification for believing it. But your knowledge that on many occasions on which you have written the list and compared it to the second, you have discovered that you have misremembered at least one capital city, provides you with excellent inductive justification for your addition. So rationality seems to require that you believe your addition and thus require that you have beliefs that you *can see* are inconsistent. This certainly seems to count as a living and everyday example.

5 A third version of the preface paradox: fallibility

However, Olin has a reply available against this case that we should consider. This is one of her responses to a more generalized version of the paradox. This version goes as follows. In the past you have discovered errors in your contingent beliefs that were, at the time, perfectly justified on the basis of memory, testimony, induction, perceptual experience and so on. You therefore have good inductive reason to think that this is true of your present corpus of beliefs. Yet each of these beliefs may be perfectly justified on the basis of independent evidence for each. Thus again you appear to have justified beliefs that are inconsistent, since if each of your beliefs is true then it is false that your present corpus of beliefs contains error. Call this version '*Fallibility*'.

Against this argument Olin claims that in order to reach its conclusion, you must reason as follows.

- (1) In the past, the body of my beliefs each justified by a method *M* of its formation (perception, memory, induction and so on) frequently contained error.
- So (2) my present body of beliefs formed by *M* contains error.

- (3) My present beliefs justified by M are $B_1, B_2 \dots B_n$
 So (4) $\sim(B_1 \& B_2 \& \dots B_n)$.

Olin observes that for this argument to go through, you must be justified in believing (3). She remarks as follows.

The fact is, of course, that most of us do not have a justified belief corresponding to (3); indeed it is arguable that none of us does. So the argument cannot be used to show that it is *in fact* rational for each of us to believe an inconsistency. However, this by no means deprives the argument of its philosophical interest. For it seems possible for a rational being (perhaps an ideally rational being) to have such a belief. So the argument may succeed in establishing that it is possible for inconsistent beliefs to be justified (2003, 67).

Perhaps Olin is correct that this would not deprive the argument of *any* philosophical interest. But surely it would deprive it of a great deal. Accepting that an ideally rational being may have justified inconsistent beliefs should engage us far less than accepting that rationality may require us as human thinkers to have inconsistent beliefs in realistic cases. In contrast, it seems that you could know what all your 196 beliefs are in *World Capitals*.

Instead, Olin objects that this inductive inference from (1) and (2) and (3) to (4) is illegitimate. Suppose that I arrive at my beliefs by reading tea leaves. That method has led to error, yet this is no reason to think that new beliefs so formed are likely to be in error, because the set of beliefs that are the negations of the former beliefs would not be less likely to contain error (Olin 2003, p. 68).

But this is true because tea-leaf reading is neither a reliable guide to truth nor to falsehood, unlike methods such as perception, memory, testimony and induction. So the likelihood of error is great in either set, but this does not undermine the inductive inference.

Olin adds that ‘Imperfectly reliable evidence entails the *absence* of an airtight connection with truth; but it obviously does not follow from this that there is a reliable, projectible connection between such evidence and *falsity*’ (2003, p. 69).

This is true but irrelevant, since all the argument needs is inductive inference from the past absence of an airtight connection of methods such as perception and memory with truth, with resulting likelihood of some error, to the *continuing absence* of this connection.

More importantly, Olin argues that accepting the possibility of justified inconsistent beliefs comes with the unacceptable consequence that you may have a pair of justified beliefs that are in explicit contradiction. As we saw in Sect. 2, that consequence is indeed unacceptable. Olin appeals to several principles to support her argument that this is the price of justified inconsistent beliefs (Olin 2003, pp. 83–84). One is the *Conjunction Principle* (CP):

If you are justified in believing that p and you are justified in believing that q , then you are justified in believing that $(p \text{ and } q)$.

For example, if you are justified in believing that p , justified in believing that q and justified in believing that $\text{not-}(p \& q)$, then you are justified in believing that p and

justified in believing that q , so by CP, you are justified in believing that $(p \ \& \ q)$. But you are also justified in believing that not- $(p \ \& \ q)$.

Another principle that may be invoked is the *Deductive Closure Principle* (DCP):

If you are justified in believing that $(p \ \& \ q)$ and you recognize that this conjunction entails that r , then you are justified in believing that r .

Now her argument goes as follows. Let us apply DCP to an inconsistent set of contents $p_1, p_2 \dots p_n$, each of which you are justified in believing. *Ex hypothesis*, you are justified in believing p_1 . But since the set is inconsistent, the conjunction $(p_2 \ \& \dots \ p_n)$ entails that not- p_1 . If you recognize this and are justified in believing this conjunction, then by DCP, you are also justified in believing that not- p_1 .

It might be proposed that DCP should be rejected in favour of the *Weak Deductive Closure Principle* (WDCP)

If you are justified in believing that p and you recognize that this entails that q , then you are justified in believing that q .

This proposal does not evade Olin's argument, since combining WDCP with CP effectively yields DCP. Suppose that you are justified in believing each of $p_1, p_2 \dots p_n$. Then by repeated applications of CP, you are justified in believing the conjunction $(p_1 \ \& \ p_2 \ \& \dots \ p_n)$ and so by DCP, you are justified in believing whatever you recognize is entailed by this conjunction.

This objection counts against *World Capitals* as well. Suppose that you are justified in believing that the capital of Afghanistan is Kabul, that the capital of Albania is Tirana, that the capital of Algeria is Algiers, ... and that the capital of Zimbabwe is Harare and that at least one of these 196 answers is false. Then you should recognize that the set of what you believe minus, say, 'the capital of Afghanistan is Kabul' entails that the capital of Afghanistan is not Kabul. So by DCP you are justified in believing that the capital of Afghanistan is not Kabul—although you are still justified in believing that it is.

6 A fourth version of the preface paradox: Modesty

Those of us who sense no irrationality in *World Capitals* might argue that DCP or CP is false.¹⁵ Rather than embarking on that discussion, let us instead consider the following 'argument from *Modesty*'. Suppose that I know that I have very often discovered that very many of my vastly large sets of my past beliefs (whether or not these beliefs are justified) include at least one false belief. This is an excellent inductive reason to believe that my present vastly large set of beliefs (whether or not these beliefs are justified) includes at least one false belief. Accordingly, I form the belief that my total beliefs—excluding this one—contain at least one false belief. By 'my total beliefs' I rigidly designate my actual beliefs. I am not thinking of false beliefs that I *might* have. If we call my 'other' beliefs $B_1 \ \& \ B_2 \ \& \dots \ B_n$, then my set of total beliefs is as follows.

¹⁵ Olin (2003, pp. 99–101) herself raises an objection to DCP.

Modesty: My beliefs (excluding *Modesty*) contain at least one false belief
&
 $B_1, B_2 \dots B_n$

As noted in Sect. 2, this is a case of exclusively self-referential inconsistency in belief. My total set of beliefs is inconsistent, because if *Modesty* is true then at least one of $B_1, B_2 \dots B_n$ must be false.

In passing, it is worth noting that *Modesty* is neither true if false nor false if true. For if *Modesty* is false, then all my total beliefs except *Modesty* are true. In that case all that follows is that *Modesty* is false and $B_1, B_2 \dots B_n$ are all true. On the other hand, if *Modesty* is true, then it does not follow that it is false, because my only false belief might be one of $B_1, B_2 \dots B_n$.

Since I enjoy excellent inductive evidence for *Modesty*, I should believe it. So rationality *requires* that I have inconsistent beliefs. Indeed, this requirement would be unchanged by my *recognition* that my beliefs are inconsistent, since that would do nothing at all to change the balance of evidence. Thus rationality may require one to hold inconsistent beliefs even while recognizing the inconsistency.

This new version evades Olin's objection to the inductive inference found in *Fallibility*. This is because it makes no appeal to methods of justification for my beliefs. All it needs is induction from the fact that many of my vastly large sets of my past beliefs (whether or not these beliefs are justified) include at least one that is false, to the conclusion that my vastly large set of present beliefs (whether or not these beliefs are justified) includes at least one that is false.

The argument from *Modesty* is also immune from Olin's objection that you get saddled with a pair of justified beliefs that are in explicit contradiction. This is because—even DCP is true—there is an obstacle to its application that seems insurmountable. To apply it to the contents of my total beliefs, one must suppose that one of these is subtracted. Then I must *recognize* that the remaining conjunction entails the negation of the subtracted conjunct. To perform this feat of recognition I must know what the remaining conjunction is. But the remaining conjunction is the conjunction of all except one of my total beliefs. Surely I cannot be reasonably expected to recognize what *all* but one of my beliefs are. As we noted in Sect. 2, many of these are unconscious, dispositional and include fleeting perceptual beliefs. Moreover, in *recognizing* what all but one of my beliefs are, I would thereby form a *belief* about what my beliefs are, which would itself have to be recognized and so on ad infinitum. Finally, in recognizing what all but one of my beliefs are, I would have to form a belief the content of which conjoins the contents of all but one my other beliefs. Surely the sheer size of this conjunction means that I cannot think the thought of it. So Searle's principle prohibits me from believing it. Acknowledging this limitation in human belief is entirely appropriate, because the argument from *Modesty* starts by acknowledging human fallibility.

It seems plausible that my inductive evidence for *Modesty* does not count against any of $B_1 \& B_2 \& \dots B_n$. Evidence for each of B_1, B_2, \dots, B_n will include evidence of every type, but evidence for *Modesty* might simply be strong induction.

The argument from *Modesty* also evades two other general objections that Olin makes against traditional versions of the paradox. The first of these is that if inconsistent

beliefs may be rational, then ‘any coherence theory of justification is misguided, since coherence presupposes consistency’ (2003, p. 62). The second is that if we accept the possibility of rational inconsistent beliefs then

... it seems that one might acknowledge the validity of a particular *reductio* while insisting that it need have no impact on belief. A properly executed *reductio* will have no epistemic force (2003, p. 86)¹⁶

Notice first of all that both objections are arguments against the *conclusion* of a supposedly paradoxical argument, in particular, the argument from *Modesty*. This means that even these arguments turn out to be apparently impeccable, they only result in sharpening the paradox, since they say nothing about the premises of the argument for the possibility of rational inconsistent beliefs. The result is that we still have apparently impeccable reasons to accept the possibility and now, apparently impeccable reasons against it. The paradox—if it still deserves to be called one—far from being dispelled, is now more entrenched. Let us examine the first objection. To elaborate on it slightly, this is that any coherence theory of epistemic justification holds that one’s beliefs are epistemically justified to the extent that they ‘cohere’ with the rest of one’s beliefs. But whatever else coherence amounts to, it requires consistency (Ewing 1934; Cherniak 1984; Lewis 1946). Thus any coherence theory of epistemic justification must say that one’s belief is epistemically justified only if it is consistent with the rest of one’s beliefs. So if inconsistent beliefs may be rational or justified, then any coherence theory of epistemic justification is false.

However, not all coherence theories elucidate coherence in terms of consistency. For example, Paul Thagard (2000) holds that contradictory propositions are incoherent with each other, but is silent about the coherence of an inconsistent set of propositions.

Secondly, a coherentist might agree that a version of the paradox shows that rational inconsistent beliefs are possible but respond that consistency is not an essential component of coherence. BonJour (1989) is a case in point.

Thirdly, some of us might be already impressed with the objections against coherence theories of epistemic justification.¹⁷ If so, then we might see Olin’s argument as terminating in the right conclusion and thus might not take it as an objection to the claim that inconsistent beliefs may be epistemically rational.

Finally, although in believing *Modesty*, I must have at least one false belief, this does not mean that my other beliefs B_1, B_2, \dots, B_n are inconsistent. So I may still sensibly aim for coherence, and thus consistency, in my other beliefs.

Now let us turn to the second objection. Many, and perhaps most, *reductios* proceed by arguing that a proposition is false because one may derive from the supposition that it is true a contradiction rather than an inconsistent set of propositions. So the impact of the possibility of rational inconsistent beliefs upon belief-revision is less

¹⁶ Klein (1985, p. 125) acknowledges this objection more tersely.

¹⁷ These include the *isolation objection*, namely that the fact that a system of beliefs is coherent cannot provide guidance to truth and reality (Moser 1989; Plantinga 1993, pp. 82–84), and its variant, the *alternative systems objection*, namely that for each coherent system of beliefs there are other incompatible possible systems that are equally coherent. Then there are the *impossibility results* that seem to show that no measure of coherence is truth-conducive (Bovens and Hartmann 2003; Olsson 2005).

than might appear. Still, there could be such *reductios*.¹⁸ But it is difficult to see how the inconsistency involved in believing *Modesty* may figure in this procedure. We would have start with a target supposition to be shown to be false and derive from it the consequence that one must have at least one false belief (excluding this). This consequence is of course possibly true and so does not fit the bill.

My belief in *Modesty* has epistemic force because it motivates me to re-examine my beliefs in search of falsehood. Indeed, it could only motivate me if I *recognize* the inconsistency. This is an instrumental intellectual good stemming from the virtue of intellectual humility in acknowledging the actualization of my own fallibility.

In sum, the argument from *Modesty* shows that rationality may require you to hold inconsistent beliefs even when you recognize that they are inconsistent. That conclusion is benign because rationality still forbids you from having explicitly contradictory beliefs.

Similarly, for Makinson's original version to succeed, my other beliefs must all be justified. Once again, the obvious strategy for meeting this requirement is to suppose that they are consistent, otherwise evidence for some will be evidence against others. Perhaps it is unlikely that any of us are consistent in our beliefs. Nonetheless, it seems to be a recognizable human ideal that we be so. Moreover, the question of whether we can be justified in all of our 'other' beliefs seems to be an empirical one. That question is not really part of the paradox, since that arises only when we add some other belief about them. Finally, it certainly seems that my belief that my total beliefs—excluding this one—contain at least one false belief counts as a 'living and everyday example'.

At this point I have achieved my objective. Nonetheless further exploration might be interesting. While I cannot claim that there are no other versions of the paradox besides the four that I have discussed, there is a loose sense in which I have exhausted them; in comparison with Makinson's original version, *World Capitals* supposes that one has fewer beliefs, Olin's *Fallibility* supposes that one has more and *Modesty* supposes that one has a belief about *all but one* of one's beliefs.

Nonetheless, my discussion would be more thorough were it to explore a fifth version, one that turns *Modesty* into a case of 'inclusive' self-reference in which one has a belief about *all* of one's beliefs Accordingly I now turn to a tentative examination of this—tentative partly because I will look at this version only with the classical assumptions that every proposition is true or false and that no proposition is both true and false.

7 A fifth version of the preface paradox: Modesty*

Let us suppose that the case of *Modesty* is exactly as described above except that this time I form the belief that my total beliefs—including this one—contain at least one false belief. My set of total beliefs is as follows.

*Modesty**: My beliefs (including *Modesty**) contain at least one false belief
&
 $B_1, B_2 \dots B_n$

¹⁸ A case in point is Hájek's type 2 Cable Guy paradox (2005).

Unlike *Modesty*, I have a belief about *all* my beliefs. If *Modesty** is false then all my beliefs are true, including *Modesty**. Believing *Modesty** guarantees that its content is true, in which case my beliefs are inconsistent, because at least one of them must be false. This result may be generalized to any believer. Once you think you're wrong, you must be right! Since I enjoy excellent inductive evidence for *Modesty**, I should believe it. So rationality requires that I have inconsistent beliefs. Indeed, this requirement would be unchanged by my recognition that my belief in *Modesty** is self-verifying, since that would do nothing at all to change the balance of evidence. Thus rationality may require one to hold inconsistent beliefs even while recognizing the inconsistency.

It is worth noting that although *Modesty** is true if false, it is not false if true. For one of my other beliefs $B_1, B_2 \dots B_n$, say B_1 , might be false, in which case *Modesty** would still be true. Supposing that there is another false belief, say B_2 , among my other beliefs $B_1, B_2 \dots B_n$ will not turn *Modesty** into a falsehood.

This version of the paradox evades all the objections that we have considered so far against the argument from *Modesty*. Moreover, it is immune in a second way from Olin's argument that it entails a pair of justified beliefs that are in explicit contradiction. To appreciate this point, we should first consider an objection that she makes.

Olin considers the second step in *Fallibility*, namely,

(2) My present body of beliefs formed by *M* contains error.

She remarks as follows.

... if all my beliefs other than (2) were true, then (2) would be true if and only if it were false. But there is paradox enough without having to worry about self-reference. My way of avoiding these problems is to use the phrase "the body of my other reasonable beliefs", where this phrase is understood to exclude each statement in the argument (Olin 2003, p. 67, note 9).

Transposing her case into mine, her argument appears to be as follows. Hypothesize that my beliefs $B_1, B_2 \dots B_n$ are all true. If *Modesty** is true then my beliefs (including *Modesty**) contain at least one false belief. *Ex hypothesis*, this false belief cannot be among $B_1, B_2 \dots B_n$. So it must be *Modesty** that is false. Thus if *Modesty** is true then it is false. And still on the hypothesis that $B_1, B_2 \dots B_n$ are all true, if *Modesty** is false, then none of my beliefs (including *Modesty**) are false. So all my beliefs, including *Modesty**, are true. Thus if *Modesty** is false then it is true. So *Modesty** is true just in case it is false. (In passing let us call this 'Olin's point'). We now have a paradox of self-reference that we should avoid. Avoiding it requires that we stipulate that *Modesty** is not included in my other beliefs.

Olin's point is correct. But an equivalent way to look at it is that because *Modesty** is true just in case it is false, then (assuming that it is either true or false) it is both true and false. Given that no proposition can be true and false, this is impossible. That impossibility is entailed by the hypothesis that $B_1, B_2 \dots B_n$ are all true. So $B_1, B_2 \dots B_n$ cannot all be true after all. In that case *once I believe Modesty**, it must be true, and so I have inconsistent beliefs. This is essentially the same result as mine!

However, the objection might be developed.¹⁹ This time it goes as follows. Suppose once again that my set of total beliefs is as follows.

*Modesty**: My beliefs (including *Modesty**) contain at least one false belief
&
 $B_1, B_2 \dots B_n$

Hypothesize that my beliefs $B_1, B_2 \dots B_n$ are all true. As shown above, it follows that *Modesty** is true just in case it is false. Since every proposition is either true or false, *Modesty** is both true and false. But no proposition can be true and false. So there is no proposition expressed by *Modesty**.

However, on the assumption (one we have made all along) that the contents of beliefs are propositions, this has the startling result that I cannot believe *Modesty** after all. Indeed, I could not even *wonder* whether it is true, given that wondering is also a propositional attitude. Of course, intuitions may vary. But it seems to me that I could not only wonder whether it is true but also believe that it is. Surely I could believe that I have at least one false belief. But I should also see that this is not a certainty, so I should acknowledge the possibility that my second-order belief, although justified, is false. Why couldn't I acknowledge this by modifying the content of that belief to *Modesty**? Doesn't 'At least one of my beliefs (including this) is false' express a clear thought?

Looking at the further-developed objection this way, what entails the *wrong* result that I cannot believe *Modesty**? Given that every proposition is either true or false and that no proposition can be both true and false, it is the hypothesis that my beliefs $B_1, B_2 \dots B_n$ are all true. So that hypothesis must be false. So $B_1, B_2 \dots B_n$ cannot all be true after all. In that case *once I believe Modesty**, it must be true, and so I have inconsistent beliefs. Once again, this is essentially the same result as mine.

There is second way of reacting to the further-developed objection. This is to argue that *if* every proposition is either true or false, *Modesty** is both true and false. But no proposition can be true and false, so *Modesty** is neither true nor false. However, even from the perspective of a truth-gap theorist, this would have the odd consequence that my attitude to *Modesty** is that it is true, although it is not true, something I could work out for myself. Assuming that this consequence is unacceptable, the upshot again is that I cannot believe *Modesty**.

This is same wrong result as before. What entails it? Given that no proposition can be both true and false, it is the hypothesis that my beliefs $B_1, B_2 \dots B_n$ are all true. So that hypothesis must be false. So $B_1, B_2 \dots B_n$ cannot all be true after all. In that case *once I believe Modesty**, it must be true, and so I have inconsistent beliefs. Yet again, this is essentially the same result as mine.

Those who are still uncomfortable with my conclusion will naturally fasten upon self-reference as an object of suspicion. They will tend to still think of the case as a paradox, and so might accuse me of solving one paradox with another. But that might not be such a bad thing, especially if the other is tractable. Moreover, if it really is a paradox then we should expect that any solution of it will be a bit counterintuitive.

¹⁹ I thank the first anonymous reviewer for making me see this.

But in any case, although self-reference is involved in my argument from *Modesty**, we have yet no reason to think that this is paradoxical or illegitimate.

First of all, some forms of self-reference are legitimate, as this very sentence shows. In addition, consider the following.

Well-evidenced: I have at least one well-evidenced belief.

If I believe *Well-evidenced*, then its content partly refers to itself, because the candidates for a well-evidenced belief include *Well-evidenced*. Likewise, if I believe *Modesty**, then its content partly refers to itself, because the candidates for a false belief include *Modesty**. Thus there is an important structural similarity between *Modesty** and *Well-evidenced* in terms of self-reference. But there is nothing illegitimate in believing *Well-evidenced*. I might sensibly reflect on some of my beliefs and recognize that one of them is well-evidenced, thus justifying my belief in *Well-evidenced*.

Perhaps the worry is that the argument trades upon the same kind of self-reference found in the Liar paradox. However, the logic of *Modesty** is different. Consider

Liar: This proposition is false.

If *Liar* is true then it says truly that it is false, hence it is false. If it is false then that is what it says, hence it is true. In contrast, the truth of *Modesty** does not entail its falsehood, because my only false belief might be one of $B_1, B_2 \dots B_n$. And provided I do not believe *Modesty**, its falsehood does not entail its truth. All that follows is that at least one of $B_1, B_2 \dots B_n$ is false.²⁰

If *Modesty** is false then my belief in *Modesty** is true. As we just saw however, it is not the case that if *Modesty** is true then my belief in *Modesty** is false, because my only false belief might be one of $B_1, B_2 \dots B_n$.

Moreover, the self-reference involved in my argument from *Modesty** differs from binary self-reference. Consider the following.

First: The next sentence in this paper is true.

Next: The previous sentence in this paper is false.

While neither of these refers to itself, each refers to the other. Thus insofar as paradoxical reference goes, it is only their *conjunction* that is of interest. In contrast, *Modesty** refers to itself. A second difference is that although *Modesty** refers to my other beliefs $B_1, B_2 \dots B_n$, none of these refer to *Modesty**. Thus strictly speaking, no binary reference is involved.

A third difference is that unlike the conjunction of the contents of my total beliefs in the case of *Modesty**, the conjunction of *First* and *Next* is true if false. For if that conjunction is false then either *First* is false or *Next* is false. But if *First* is false then *Next* is false (since *First* says falsely that *Next* is true) and hence *First* is true (since *Next* says falsely that *First* is false). On the other hand, if *Next* is false then *First* is true (since *Next* says falsely that *First* is false) and hence *Next* is true (since *First* says truly that *Next* is true). So the conjunction of *First* and *Next* is true.

²⁰ It might be objected that *Modesty** presupposes that I believe it, perhaps in the sense that *Modesty** lacks a truth-value if I do not believe it (see Strawson 1952, p. 216). This snag may be avoided by supposing that we read *Modesty** as ‘My total actual beliefs contain at least one false belief’.

In contrast, the conjunction of the contents of my beliefs in the case of *Modesty** might be false because my only false belief is B_1 . In that case the conjunction of the contents of my total beliefs remains false (although *Modesty** is true).

We are now in a position to see that my argument from *Modesty** is immune in a second way from Olin's objection that it saddles you with a pair of justified beliefs that are in explicit contradiction. This is that CP cannot be applied to my argument. Recall how CP was used in the case in which you are justified in believing each of p , q and not- $(p \ \& \ q)$. CP delivers the desired result because it is applied to a set of beliefs, one of which, i.e. not- $(p \ \& \ q)$, is in explicit contradiction with the conjunction of the others in that set, i.e. p , q . The contradiction is syntactic, so that one might be expected to recognize the contradiction. More importantly, the contradiction is straightforwardly *classical*. In other words, although exactly one of ' $(p \ \& \ q)$ ' and 'not- $(p \ \& \ q)$ ' must be true, neither is supposed to be both true *and* false. The point of applying CP to your set of contents [p , q and not- $(p \ \& \ q)$] is that if you have justification for each, then you will end up with justification for believing that $(p \ \& \ q)$ and with justification for believing that not- $(p \ \& \ q)$. This is impossible, because justification for the truth of either belief will be justification for the falsehood of the other.

The same cannot be said of *Modesty**. For the conjunction of my other beliefs, namely $(B_1 \ \& \ B_2 \ \& \ \dots \ B_n)$ is *not* in explicit *classical* contradiction with *Modesty**. For Olin's point is that once I believe *Modesty**, the truth of the conjunction of my other beliefs, namely $B_1 \ \& \ B_2 \ \& \ \dots \ B_n$, entails that *Modesty** is both true *and* false. By CP, I have justification to believe the conjunction $B_1 \ \& \ B_2 \ \& \ \dots \ B_n$. But is my justification for believing this conjunction justification for believing that *Modesty** is false? If it provides me with any justification at all for believing anything about the truth-value of *Modesty**, then it is justification for believing that *Modesty** is both true *and* false. Given that this is impossible, at least from a classical perspective, this is not justification for believing anything about *Modesty**.

8 Concluding remarks

To recap, I started naturally enough by examining Makinson's original version of the paradox, but concluded that it does not seem to be a living and everyday example. *World Capitals* avoids this difficulty but is vulnerable to Olin's objection that accepting the possibility of justified inconsistent beliefs saddles you with a pair of justified beliefs that are in explicit contradiction. In contrast, *Modesty*—and arguably *Modesty** as well—escapes all the objections that could trouble the other versions. These are living and everyday cases in which rationality requires you to have inconsistent beliefs even while you recognize that your beliefs are inconsistent.²¹

References

Audi, R. (1994). Dispositional beliefs and dispositions to believe. *Noûs*, 28(4), 419–434.

²¹ I am grateful to three anonymous reviewers whose valuable insights compelled me to radically rethink drafts of this paper in many ways.

- Barnes H. E. 1958. (Trans.). J. Sartre. *Being and nothingness*. London: Methuen.
- BonJour, L. (1989). Replies and clarifications. In John W. Bender (Ed.), *The current state of the coherence theory* (pp. 276–292). Dordrecht: Kluwer.
- Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Clarendon Press.
- Burge, T. (1978). Belief and synonymy. *Journal of Philosophy*, 75(3), 138–199.
- Cherniak, C. (1984). Computational complexity and the universal acceptance of logic. *Journal of Philosophy*, 81(12), 739–758.
- Chisholm, R. (1989). *Theory of knowledge* (3rd ed.). Englewoods Cliffs, NJ: Prentice Hall.
- de Almeida, C. (2001). What Moore's paradox is about. *Philosophy and Phenomenological Research*, 62(1), 33–58.
- de Almeida, C. (2012). Epistemic closure skepticism and defeasibility. *Synthese*, 188(2), 197–215.
- Ellis, B. (1979). *Rational belief systems*. Oxford: Basil Blackwell.
- Ewing, A. C. (1934). *Idealism: A critical survey*. London: Methuen.
- Foley, R. (1979). Justified inconsistent beliefs. *American Philosophical Quarterly*, 16(4), 247–257.
- Foley, R. (1992). Being knowingly incoherent. *Noûs*, 26(2), 181–203.
- Hájek, A. (2005). The cable guy paradox. *Analysis*, 65(2), 112–119.
- Hempel, C. (1962). Deductive-nomological vs. statistical explanation. In H. Feigl & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science* (Vol. 3, pp. 69–98). Minneapolis: University of Minnesota Press.
- Hintikka, J. (1962). *Knowledge and belief*. Ithaca, NY: Cornell University Press.
- Hoffman, R. (1968). Mr. Makinson's paradox. *Mind*, 77(305), 122–123.
- Klein, P. (1985). The virtues of inconsistency. *The Monist*, 68(1), 105–135.
- Lacey, A. (1970). The paradox of the preface. *Mind*, 79(316), 614–615.
- Lehrer, K. (1974). *Knowledge*. Oxford: Clarendon Press.
- Leitgeb, H. (2014). A way out of the preface paradox? *Analysis*, 74(1), 11–15.
- Lewis, C. I. (1946). *An analysis of knowledge and valuation*. LaSalle: Open Court.
- Makinson, D. (1965). The paradox of the preface. *Analysis*, 25(6), 205–207.
- Moser, P. K. (1989). Lehrer's coherentism and the isolation objection. In J. W. Bender (Ed.), *The current state of the coherence theory* (pp. 29–37). Dordrecht: Kluwer.
- New, C. (1978). A note on the paradox of the preface. *The Philosophical Quarterly*, 28(113), 341–344.
- Olin, D. (2003). *Paradox*. Montreal: McGill-Queen's University Press.
- Olsson, E. J. (2005). *Against coherence: Truth, probability, and justification*. Oxford: Clarendon Press.
- Passmore, J. (1957). *A hundred years of philosophy*. London: Duckworth.
- Plantinga, A. (1993). *Warrant: The current debate*. New York: Oxford University Press.
- Pollock, J. (1983). Epistemology and probability. *Synthese*, 55(2), 231–252.
- Pruss, A. (2011). Does belief distribute over conjunction? <http://alexanderpruss.blogspot.sg/2011/10/does-belief-distribute-over-conjunction.html>.
- Quine, W., & Ullian, J. (1970). *The Web of belief*. New York: Random House.
- Routley, R., & Routley, V. (1975). The role of inconsistent and incomplete theories in the logic of belief. *Communication and Cognition*, 8, 185–235.
- Ryan, S. (1991). The preface paradox. *Philosophical Studies*, 64(3), 293–307.
- Sainsbury, R. M. (1995). *Paradoxes* (2nd ed.). Cambridge: Cambridge University Press.
- Schick, F. (1963). Consistency and rationality. *The Journal of Philosophy*, 60(1), 5–19.
- Searle, J. (1992). *The rediscovery of the mind*. Cambridge, MA: MIT Press.
- Sorensen, R. A. (1985). Pure Moorean propositions. *Canadian Journal of Philosophy*, 15(3), 489–505.
- Stalnaker, R. (1987). *Inquiry*. Cambridge: MIT Press.
- Strawson, P. F. (1952). *Introduction to logical theory*. London: Methuen.
- Swain, M. (1970). The consistency of rational belief. In M. Swain (Ed.), *Induction, acceptance and rational belief* (pp. 27–54). Dordrecht: Reidel.
- Thagard, P. (2000). *Coherence in thought and action*. Cambridge, MA: The MIT Press.
- Williams, J. N. (2007). Moore's paradoxes and iterated belief. *Journal of Philosophical Research*, 32, 145–168.
- Williams, J. N. (2014). Moore's paradox in belief and desire. *Acta Analytica*, 29(1), 1–23.