Journal of the American Philosophical Association (2023) 154–171
DOI:10.1017/apa.2021.48

What are Paradoxes?

ABSTRACT: *According to a standard view, paradoxes are arguments with plausible premises that entail an implausible conclusion. This is false. In many paradoxes the premises are not plausible* precisely because *they entail an implausible conclusion. Obvious responses to this problem—including that the premises are individually plausible and that they are plausible setting aside the fact that they entail an implausible conclusion—are shown to be inadequate. A very different view of paradox is then introduced. This is a functionalist view according to which paradoxes are the kinds of things that puzzle people in characteristic ways. It is claimed that this view, too, fails and for the very same reason. The result is a new puzzle about the nature of paradoxes.*

KEY WORDS: paradoxes, methodology, logic

Philosophy is full of paradoxes, but what exactly are they? One might expect to find a significant body of work in answer to this question. Curiously though, existing accounts are few, far between, and surprisingly problematic. The aim of this paper is to illustrate how deep these problems run, to motivate an alternative view of what paradoxes are and, ultimately, to generate a new—and I hope interesting—puzzle about their nature. My core claims are as follows. First, I claim that orthodox views of paradox are vulnerable to a challenge that I refer to as the problem of paradoxical involvement. Second, there is a curiously neglected alternative to the orthodoxy, which I refer to as a functionalist view of paradox. Third, I claim that while functionalist views are initially promising, they ultimately face the very same challenge as orthodox views. Fourth, the result of the three foregoing points—fittingly enough for a paper about paradoxes—is a puzzle: we may know a paradox when we see one, but we do not, at present, know what it is to be one.

Before beginning, a methodological point is in order. I assume that there is an account of paradox to be found and that it is fairly well-behaved. I do not, for example, discuss 'family resemblance' views or messy disjunctive views. This assumption may be incorrect. Indeed this paper may be evidence that it *is* incorrect. Nonetheless, it is my working hypothesis. I invite the reader to share it for now and see where they end up.

I am grateful to Durham's Mind, Language and Metaphysics Research Cluster for the chance to present a draft version of this paper. Thanks also to David Faraci and Louise Hanson (to whom I owe the label 'functionalist') for early discussion and to two anonymous reviewers and the associate editor at the *Journal of the APA*.

1. Preliminaries: Between Subjective and Objective

Consider the following first-pass from the introduction to Mark Sainsbury's book, *Paradoxes*:

> This is what I understand by a paradox: an apparently unacceptable conclusion derived by apparently acceptable reasoning from apparently acceptable premises.
>
> (1988: 1)

This will likely ring true to many philosophers. Think about the paradox of the heap. You have a heap of stones. Taking one away will not make it not-a-heap. So you take one away and you have still got a heap. Continue to do this until you have one stone left. By the above reasoning you have still got a heap. But this must be false. A single stone is not a heap. This looks exactly as Sainsbury describes it: an apparently unacceptable conclusion derived from apparently acceptable premises and apparently acceptable reasoning. Or consider the lottery paradox. Your ticket has a low enough probability of winning that it is reasonable to believe it will not win. The same is true of all of the other tickets. It is reasonable, then, to believe of each ticket that it will not win. If it is reasonable to believe of each that it will not win, then it is reasonable to believe of all that none will win. Thus, it is reasonable to believe that none will win. But this is false. You know, assuming all tickets have been sold, that one will win. Again, this looks like a case in which an apparently unacceptable conclusion is derived from apparently acceptable premises and apparently acceptable reasoning.

There is a problem. It concerns the use of what we might call a subjective qualifier in the foregoing account of paradox. In Sainsbury's account that qualifier is *appearance*. Appearance qualifies the acceptability of the premises, the acceptability of the reasoning, and the acceptability of the conclusion. This leaves the account incomplete. Things do not just appear some way or other. They always appear that way to someone. To whom are the appearances that constitute a paradox doing their appearing? The question is not for Sainsbury (1988) alone; it applies equally to many variants on his view, scattered through the literature. For example, in discussing axiological paradoxes Larry Temkin writes: 'I believe that a position is paradoxical if it involves two or more incompatible views, each of which seems, even on reflection, intuitively obvious, certain or (virtually) undeniable' (2012: 298).Temkin's account differs from Sainsbury's in many respects. Note though that, like Sainsbury, Temkin also makes use of a subjective qualifier: seems. The incompatible views *seem* obvious. But seems to whom? Again, without an answer to this question the account is incomplete. Or consider Quine: 'A conclusion that at first sounds absurd but that has an argument to sustain it' (1962:84). Sounds to whom? Insofar as the need to answer this question has been recognized in the literature it clearly has not been regarded as posing much of a problem. In his book *Paradoxes from A to Z*, Michael Clark (2012: 160) writes:

> But what you might ask counts as acceptable and unacceptable? (Un) acceptable to whom? Isn't Sainsbury's account too vague? No, on the

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

> contrary, the vagueness . . . is an advantage since what counts as contrary to received opinion will vary with that opinion.

Clark is clearly addressing the kind of problem—with subjective qualifiers—that I raised above. But he does not think it is a problem. His thought seems to be that 'what counts as acceptable and unacceptable' will vary with opinion. The obvious implication to draw from this is that what counts as a paradox will vary with opinion too. Clark thinks that is fine.

Let us dig a bit deeper. What exactly does it amount to if we allow something's status as a paradox to 'vary with opinion'? It would appear to lead toward a view of paradoxes as in some way relative to inquirers. We find this point made quite explicitly by the only other author to consider the role of subjective qualifiers in accounts of paradox. In his *Brief History of the Paradox*, following a discussion of the seventeenth-century mathematician and probability theorist Gerolamo Cardano's work on the behavior of dice, Roy Sorensen suggests that: '"Paradox" should be relativised to the thinker in question' (2003: 224). Sorensen is going slightly further than Clark, explicitly favoring accounts of paradox that are in some way 'relative'. How exactly should this relativity be understood? One simple view would be to understand 'paradox' as describing a two-place relation: one place for an argument and one for a person to whom the argument's features do their appearing or seeming. On this view utterances of the form 'this is a paradox' would vary with the views of the speaker roughly as follows:

> (A) For any argument *A* and any person S: If *A* has premises that seem to S be true and that entail a conclusion that seems to S be false, then '*A* is a paradox', as stated by S, is true.

This is what we might call a strongly subjective account of paradox. Two points of housekeeping are in order before we assess it. First, (A) assumes—as I largely shall—that paradoxes are types of argument rather than types of conclusion. I do not think much turns on this. Another alternative, which I discuss below, is to understand paradoxes as sets of inconsistent propositions. Second, I do not refer to reasoning in (A); only to premises, conclusions, and the entailment relations between them. Again, I do not take anything to turn on this; any reasoning can be represented within the premises themselves.

Can (A)—or something like it—be correct? Surely not. It is far too permissive. It makes paradoxes too easy to come by. Consider Irrational Ian. Ian has a bizarre set of beliefs. He believes both that grass is green and that grass is colored, but he also believes that green is not a color. He therefore finds the following argument troubling:

1. Grass is green
2. If grass is green then grass is not colored.
3. (1, 2) Grass is not colored.

This seems to Ian to have true premises that entail a conclusion that seems to him false. If (A) were true, then when Ian says 'this is a paradox' he has said something true. This is surely too permissive. He has not said something true. Paradoxes cannot be that easy to come by; you do not get one on every occasion that a sufficiently eccentric person expresses their beliefs. There must be some constraints on the appearances that can constitute a real or genuine paradox. That, I submit, is the ordinary view at least. And that is not all. There are more generic problems with views like (A), problems that apply to crude forms of relativism, of which (A) is an instance. One concerns disagreement. Defenders of (A) will have trouble explaining how two or more agents can agree or disagree about whether an argument is a paradox. Another concerns retraction. Defenders of (A) will struggle to explain why agents can sensibly retract their earlier assertions about whether some argument is a paradox.

Surely then the correct account of paradox must be less subjective than (A). One extreme would be to abandon the subjective qualifier altogether in favor of what we might call a strongly objective account:

> (B) A paradox is an argument with premises that *are* true and that entail a conclusion that *is* false.

This solves the problem of Irrational Ian. Clearly though, (B) is an overreaction to (A)'s failings. After all, (B) seems to entail that there are no paradoxes. This is because if an argument has true premises and if those premises entail the conclusion of the argument, then the conclusion is true. And if the conclusion is true, then, assuming the law of noncontradiction holds, it is not false. But (B) says that paradoxes are arguments with true premises that entail a false conclusion. Thus, (B) seems to entail that there are not any paradoxes. This is surely false. There are lots of paradoxes. Philosophy is full of them. Therefore, (B) is false.

Perhaps this is too quick. Dialetheists do not think that the law of noncontradiction holds in an unrestricted sense (e.g., Priest 2006, Priest, Beall, and Armour-Garb 2004). Furthermore, one of their classic arguments for this is that it explains paradoxes. Therefore, (B) should not be dismissed so quickly.

Still, it should ultimately be dismissed. Even if we grant that dialetheism is true and that it can explain some paradoxes it certainly cannot explain all paradoxes. The dialetheist's strong suits are the paradoxes of set-theory and self-reference. While some dialetheists have tried to extend their approach beyond this familiar ground—for example, to the paradox of the arrow (different from Arrow's paradox!) and to the sorites paradox (e.g., Priest 2006 chapters 11–14)—there are still plenty of paradoxes that are clearly not amenable. Epistemic paradoxes—for example, the lottery paradox and the preface paradox—are good examples, as are pragmatic paradoxes (e.g., Moore's paradox) and many ethical paradoxes (e.g., the paradox of deontology, the paradox of deterrence) in which dialetheist responses would clearly fail to get purchase.

Thus, (A) is 'too subjective' and (B) is too objective. We need something in between. Consider the following accounts offered by Nicholas Rescher in his book

Paradoxes: Their Roots, Range and Resolution (2001) and by Bill Lycan (2010), respectively:

> Paradoxes thus arise when we have a plurality of theses, each individually plausible in the circumstances, but nevertheless in the aggregate constituting an inconsistent group.
>
> (2001: 7)

> A paradox is an inconsistent set of propositions each of which is very plausible.
>
> (2010: 618)

There are two points to note. These accounts refer to plausibility. This is promising given our aims. These accounts seem to ask more by way of objectivity than (A), but less than (B). They ask for more that (A) because Irrational Ian's views are not plausible, and less than (B) because plausibility does not entail truth. This is promising. Let us see if we can develop a good account of paradox on this basis, that is, an account based on plausibility.

Again, one small piece of housekeeping before we begin. Rescher and Lycan both understand paradoxes as sets of propositions. For ease of presentation however I am going to stick to my earlier view of paradoxes as arguments. Nothing in what follows turns on this. While there are undoubtedly advantages to thinking of paradoxes as sets of propositions rather than as arguments, the basic issues that I will develop in the remainder of this article could equally well be applied to either. The important point for our purposes is really just whether the appeal to plausibility provides the middle ground between (A) and (B) that we are looking for. To see whether it does, let us start with the following formulation:

> (C) A paradox is an argument with premises that it is very plausible are true and that entail a conclusion that it is very plausible is false.

This is, I think, a fairly standard or orthodox view of the nature of paradox. Does it provide the middle ground between (A) and (B) that we are looking for?

2. Assessing the Orthodoxy: Some Preliminaries

My aim is to develop a new and interesting challenge to accounts like (C). My claim *will not* be that it fails to provide *a* middle ground between (A) and (B). My claim will be that it provides the *wrong* middle ground. It provides the wrong middle ground because it is too restrictive as an account of paradox. In this section I make some important preliminary remarks that will set up this new challenge.

I am aware of just one challenge to (C) in the existing literature. It is that (C) is vulnerable to what Sorensen describes as 'jumble arguments'. Consider the last 1,000 sentences that I have uttered. I believe each one considered individually; after all, I uttered them. But I also believe that the conjunction of all 1,000 is highly likely to be inconsistent. This is for the unexciting reason that, like anyone

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

else, if I say enough I will contradict myself. Now if a conjunction is inconsistent, then anything—including obviously false propositions—are entailed by it. It follows from this and (C) that an argument consisting of the last thousand sentences I have uttered, followed by any arbitrary false conclusion is a paradox. It is a paradox because each premise is very plausibly true; yet, together these premises yield a conclusion that is obviously false.

Sorensen thinks this undermines views like (C) because jumble arguments are not paradoxes. They are just a jumble of any old individually plausible but jointly inconsistent propositions. Accordingly, (C) is too permissive; it entails that arguments that are not paradoxes are paradoxes. I will return to jumble arguments later, but for now I want to set them to one side. My aim is to develop a new objection to (C). Unlike jumble arguments, which aim to show that (C) is too permissive, I aim to show that (C) is too restrictive. I claim that (C) entails that arguments that are paradoxes are not paradoxes. This is because, I claim, there are—or could be—paradoxes that do not have very plausible premises or do not have very implausible conclusions.

On hearing that I will be arguing in this way—that is, by claiming that there are paradoxes that do not have very plausible premises or do not have very implausible conclusions—one might anticipate the following, rather flat-footed kind of argument: some claims that used to be plausible are not so anymore, but the arguments they figure in should still qualify as paradoxes. Consider as an illustration Zeno's famous paradox of the racecourse. In order to catch the tortoise, Achilles must first cover half of the distance between them. He must then cover half of the remaining distance again. He must then cover half of the remaining distance again. And so on *ad infinitum*. The need to cover half the remaining distance never comes to an end. So Achilles never catches the tortoise. Is this a paradox? Surely it is! It is the paradox *par excellence*. Yet, its premises are not very plausible. Specifically, it relies on the implausible premise that (roughly) an infinite sequence cannot sum to the finite number to which it tends. They both can and do. Accordingly, the premise on which Zeno's paradox relies is quite implausible. That is, (C) entails that the paradox *par excellence* is not a paradox, thus, not-(C). Quite generally, (C) is too restrictive because it entails that many historical paradoxes—like Zeno's—are not paradoxes because they rest on claims we now see to be implausible.

This flat-footed argument seems to fit the description of the basic concern with (C) that I raised above: (C) is too restrictive because there are paradoxes that do not have very plausible premises or do not have very implausible conclusions. It is important to be clear, however, that this is not the kind of argument I have in mind. Seeing this is important not only to avoid misunderstandings, but because it helps us to get a better handle on (C) itself and in particular to see how (C) allows us to occupy a middle ground between (A) and (B).

The reason that (C) does not (falsely) entail that the racecourse paradox is not a paradox is that the view of plausibility appealed to in the argument to the contrary is far too restrictive. The objection to (C) above relied on claiming that Zeno's paradox has an implausible premise about sums of infinite sequences. That is unnecessarily brittle. There is a perfectly respectable way of thinking about plausibility

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

according to which a proposition's plausibility is indexed to what we might call its epistemic context, that is, a context in which only certain information is available to enquirers. This helps because indexed to Zeno's epistemic context, the transfinite arithmetic needed to solve the paradoxes was unavailable. Thus, in his context, the premises (perhaps) *were* plausible.

Crucially, this more flexible way of reading 'plausible' need not be so flexible as to make (C) fall back into an anything-goes overly permissive view like (A): a view that allows that Irrational Ian speaks truly when he says of some nonsense that it is a paradox. Zeno's beliefs were plausible in his epistemic context. Ian's, we can safely assume, are not very plausibly true in his epistemic context. Therefore, (C) allows that Zeno's paradoxes are paradoxes, while Ian's so-called paradoxes are not. That is how (C) can occupy the middle ground between (A) and (B) that we were looking for. Of course, questions of detail will still remain here. While on one obvious interpretation (C) will allow that Zeno would have spoken truly were he to have said of the racecourse paradox that it is a paradox, should we interpret it so as to allow that we would speak truly in calling it a paradox? Or does it merely allow that we would speak truly in saying that it was a historical paradox, or a paradox for Zeno, or something similar? These are difficult questions. Answering them would require going into some detail regarding the sense in which the context of the uses of paradox determines its truth-conditions. There is much excellent work on this topic, but I shall set it aside for the present. It is not my concern. My reason for thinking that orthodox accounts like (C) are too restrictive is, I think, more interesting; or at least more unusual. I refer to it as the problem of paradoxical involvement.

3. Assessing the Orthodoxy: The Problem of Paradoxical Involvement

In its simplest (first) formulation, the problem for (C) is that there can be paradoxes in which the premises are not very plausible precisely because they figure in a paradox and therefore entail a highly implausible conclusion. That is, (C) is false. Imagine a paradox that has as its conclusion some proposition that is so wildly implausible that we can safely say that we know it to be false; for example, the lottery paradox in which the conclusion is clearly contradictory. Now imagine for ease that there are just two premises, p_1 and p_2, that entail—and, let us say, are known to entail—such a conclusion. There is a perfectly good sense in which we can claim to know that these premises are not very plausible. We know this precisely because they entail a proposition we know to be false. This is an objection to (C). It is an objection because (C) says that in paradoxes the premises *are* very plausible. But we have just seen that they need not be. They need not be precisely because they entail an obvious, known falsehood, as the premises of many paradoxes do. Hence the name the problem of paradoxical involvement. It is the very fact that some premises figure in a paradox that render them implausible.

To be clear then, the worry for (C) that we are considering takes roughly the following form:

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

> *The (simple) problem of paradoxical involvement*: In some paradoxes the premises are not plausible precisely because they entail a conclusion that is known to be false. This means that sometimes the premises of a paradox are not plausible. And this means that accounts of paradox that claim that in paradoxes the premises are plausible, fail.

This is the basic problem. Note that it applies equally if we think of paradoxes as sets of propositions rather than as arguments. So understood, the problem is that in some paradoxes all but one of the propositions that make up the inconsistent set are not plausible precisely because they entail the remaining member of the set, which is known to be false.

How should a defender of (C) reply? Several responses will likely have occurred to the reader already. I shall run through some of them now, explaining why they are more problematic than they at first appear.

The first response is that we should distinguish between the premises being individually very plausible and jointly very plausible. When we do this, the problem of paradoxical involvement disappears. To see this, return to my schematic example above in which p_1 and p_2 jointly entail a wildly implausible conclusion, c, which is known to be false. While this may indeed mean that p_1 and p_2 are not *jointly* very plausible it does not mean that they are not *individually* very plausible. This can be used to save (C) because (C) should—properly interpreted—state that a paradox is an argument in which the premises are *individually* very plausible.

To be clear, we are proposing responding to the problem of paradoxical involvement by adopting the following disambiguation of (C):

> (D) A paradox is an argument with premises each of which is *individually* very plausibly true and that entail a conclusion that is very plausibly false.

This not only appears to resolve the problem, but it is also true to the ordinary understanding of what a paradox is. I quoted earlier from Rescher and Lycan. Rescher (2001) explicitly refers to the propositions that make up a paradox as being individually plausible. Lycan (2010) refers to each being plausible. Sorensen (2003) is surely correct, then, in describing the ordinary understanding of a paradox as follows:

> Paradoxes have convinced many philosophers that they have a small set of beliefs that are *individually* plausible but jointly inconsistent.
>
> (2003: 107, italics mine)

It seems that (D is both a good response to the problem of paradoxical involvement and is independently motivated. Is it? I do not think so. That is, (D) is not an adequate response to the problem. I offer two arguments for this. The first is that we can imagine cases of paradox in which the fact that the premises *jointly* entail a conclusion that is known to be false *does* suffice to render each premise *individually* less than very plausible. This will happen where the conclusion is established by a small number of premises that have roughly equal degrees of plausibility and that are jointly highly implausible. Imagine, for example, a paradox in which a conclusion, *c*, that is known to be false is claimed to follow from two major premises, p_1 and p_2, where p_1 and p_2 appear to have roughly equal degrees of plausibility. This will be enough to render p_1 and p_2 jointly less than very plausible. But it may also be enough to render each of p_1 and p_2 individually less than very plausible. The reasoning is simple. Given that p_1 and p_2 jointly entail a known—and, let us suppose, obvious—falsehood, their joint plausibility is very low: let us say zero or near zero. Now add to this that, *ex hypothesi*, the plausibility of each of p_1 and p_2 is known to be roughly equal. We can conclude that the probability of *each individually* will not be significantly above zero. This is just because if two premises of the same levels of probability jointly entail a conclusion of near zero probability, then the probability of each premise individually will not be significantly above zero either. For example, let us suppose that p_1 and p_2 jointly entail *c* and that $pr(c)$ is very low, that is, $\cong 0.1$ Let us also suppose that p_1 and p_2 are of comparable probabilities, namely, $pr(p_1) \cong pr(p_2)$. The result is that $pr(p_1) \cong pr(p_2) \cong 0.3$ (given that $pr(c) = \sqrt{0.1} \cong 0.3$). That is surely too low a value to say that either p_1 or p_2 is individually very plausible. We can run the same argument with paradoxes of more than two premises, provided the conclusion that they entail is sufficiently implausible. Suppose that *c* is jointly entailed not by p_1 and p_2, but by p_1, p_2, and p_3. It will still be the case that $pr(p_1) \cong pr(p_2) \cong pr(p_3) \cong 0.3$ provided that $pr(c)$ is low enough (i.e., $\cong 0.025$). Obviously, assigning precise probabilities to the conclusions of (many) paradoxes is rather silly. But the point is clear enough. Some paradoxes have obviously false conclusions, known to be such. That—along with the possibility that the premises of those paradoxes are of roughly equal plausibility—is all we need to get the problem of paradoxical involvement up and running against (D). This argument requires three points of clarification.

The first point of clarification is that I am assuming above that plausibility is understood in terms of probability. This is not the only way of understanding plausibility, and indeed there may be some difficult cases that give pause for thought in this regard. Nevertheless, it is a natural—and, I hope, not unreasonable—assumption in most circumstances.

The second point of clarification is that I have assumed that the schematic premises (p_1, p_2, p_3) are probabilistically independent. If we relax this, the assignments to each premise could be higher than I have allowed, potentially above 0.5. For example, for a two-premise argument with a conclusion of probability 0.1 and probabilistically dependent premises of roughly equal values, the assignment to each premise could be as high as 0.55. Would this not mean that that a paradox could be an argument with individually very plausible premises and a very plausibly false conclusion after all, as (D) states? This is a sensible objection, but it does not succeed in rehabilitating (D). First, unless we assume the truth of (D), which we should not, we cannot assume that there could be no two-premise paradoxes with probabilistically independent premises. Second, even in cases in which the premises are not probabilistically independent, the

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

assignments to the premises, at least for two-premise paradoxes, could still not be significantly higher than 0.5 and 0.55 is arguably not high enough to be characterized as 'very plausible. Third, for paradoxes with conclusions that are known to be false—a class that would at least include paradoxes with contradictions as conclusions and perhaps a good many more—the probability of the conclusion is 0, and so, even if the premises were not independent, their value could be no higher than 0.5, which is surely not high enough to count as very plausible.

The third point of clarification is that the above argument is in part dependent on focusing on paradoxes with a small number of premises; I have worked with two- and three-premise paradoxes. This is because for paradoxes with larger numbers of premises (i.e., four or more), the probability of each premise could feasibly be high enough that each would be very plausible, while the conclusion would remain very implausible. One might wonder, then, exactly how many paradoxes there actually are with a small number of premises. This is an interesting question, but I shall not attempt to provide a taxonomy of paradoxes with low numbers of premises here. The fact that, as I have shown, there could be paradoxes with low numbers of premises is sufficient to achieve my goal of looking for an alternative to (D).

With these clarifications in place, I now present a second way in which we can use the problem of paradoxical involvement against (D). So far I have used the problem of paradoxical involvement to target the claim, made in (C) and (D), that paradoxes have plausible premises. But we could use much the same argument structure to target either (C) or (D)'s claim that paradoxes have implausible conclusions too. The idea is simple. Some paradoxical arguments will have conclusions that are not implausible precisely because those conclusions are entailed by premises that are both individually plausible and known to be true.

Consider the example of the twins paradox (this example may not be perfect but hopefully it illustrates the basic point effectively enough, but feel free to substitute a better example if you wish). This paradox has as its premises the core claims of special relativity. These premises, each of which is individually plausible, jointly entail the conclusion that of two twins if one leaves earth on a rocket ship travelling at an appreciable percentage of the speed of light while the other remains, the former will return to find her twin has aged more than she has. These two people born on the same day will, at the same future point in their lives, be different ages. This is a paradox—and a very famous and much-discussed one at that. But contrary to both (C) and (D) its conclusion is not very implausible. It is not very implausible precisely because we know that the premises of special relativity, which are individually very plausible, are true and that they entail the conclusion. And generally if we know some premises that are individually plausible to be true and to entail a conclusion, then the conclusion is not very implausible.

To be clear then, this second argument against (D) is making use of a more general version of the problem of paradoxical involvement, a version that allows us to target not only the claim that paradoxes have plausible premises but also the claim that they have implausible conclusions. It is as follows:

> *The (general) problem of paradoxical involvement*: In some paradoxes the premises are not plausible precisely because they entail a conclusion that is known to be false. Equally, in some paradoxes the conclusion is not implausible precisely because it is entailed by premises known to be true. In either case, accounts of paradox that claim that the premises are plausible and the conclusion is implausible, fail.

This generalized version of the problem of paradoxical involvement applies straightforwardly to both (C) and (D).

I do not think, then, that (D) is a good response to the problem of paradoxical involvement. Is there a better alternative? Perhaps the obvious option would be to add a qualification to (C) according to which in a paradox the premises are very plausible aside from the fact that they figure in the paradox. This would seem to allow us to get around the problem of paradoxical involvement fairly directly. We can represent the view as follows:

> (E) A paradox is an argument with premises that—aside from the fact that they entail a conclusion that it is very plausible is false—are individually very plausibly true and that entail a conclusion that—aside from the fact that it is entailed by premises that it is very plausible are true—is very plausibly false.

This is better than (D) but it will not work either. The issue here is slightly more delicate. It comes when we try to get a handle on what exactly it means to set aside the information that (E) is asking us to. Focus just on the premises. What exactly are we setting aside when we claim that they (i.e., the premises) are very plausible setting aside the fact that they entail a conclusion that is very implausible. There are a number of different options. Let us start with the most minimal option. We should make our assessments of p_1 and p_2 setting aside only the following propositions:

- p_1 and p_2 jointly entail *c*.
- It is very implausible that *c*.

This will not do. It is too minimal. It sets aside too little. Specifically, it does not set aside the evidence for the implausibility of the conclusion. If we do not set this evidence to one side—that is, if all the propositions that are evidence for the conclusion's implausibility are still in play when we assess the premises—then the premises could still end up being implausible. To take an admittedly crude example, but one that is helpful in making the basic point, suppose that you are assessing the premises of an argument for the conclusion that there is no external world. Your initial reaction is that the premises are bound to be implausible precisely because they entail that there is no external world, a conclusion that you know to be false. You are now told to set your knowledge that there is an external world to one side in the assessment of the premises, but you are also told not to set to one side all of the evidence in virtue of which you know that there is no

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

external world. For example, you are not setting to one side the that fact that you have hands or that there is a table in the corner of the room, and so on. Will you now find the premises plausible? Surely not. If you found the premises implausible in the first place (i.e., before setting anything to one side), then you would surely still find them implausible now. After all, they are inconsistent with your having hands!

Thus, if (E) is to work, we clearly need to work with a less minimal reading of what is set aside in appraising the premises (and conclusion) of a paradox. The obvious option is that we should set aside not only the fact that the premises entail a very implausible conclusion, but also the background evidence for this very implausible conclusion. This would fix the problem we have just encountered with the minimal reading. Suppose then that we go for a maximal reading: we set aside not only the fact that the conclusion is very implausible, but also *all* of the propositions that constitute our evidence for that conclusion. To return to our schematic example, we are setting aside:

- p_1 and p_2 jointly entail c.
- It is very implausible that c.
- The evidence for c's being implausible, that is, e_1, e_2, e_3 . . .

Unfortunately this maximal approach cannot be right either. It cannot be right because, in some cases at least, it sets aside too much. It sets aside too much because there could be cases (of paradox) in which some of the evidence for c's being implausible is also evidence or one or more of the premises, evidence without which they would not be plausible! To return to our previous example, if I set aside all of the evidence for the existence of an external world, then it is unlikely I will have enough evidence left to find the premises of this or any argument plausible!

Therefore, (E) faces a prima facie problem. The minimal reading is pretty clearly too minimal, and the maximal reading is also clearly too maximal. If (E) is to work, there must be a middle ground, some amount of information that we can set to one side that allows premises to be plausible, yet also allows the conclusion to be implausible. The issue is that it is hard to find principled reason to think there is such Goldilocks-esque middle ground. I cannot see any principled reason why there could not be a genuine paradox that consists of premises p_1, p_2, p_3 (etc.), a conclusion, c, and some item(s) of evidence, e, without which c would not be implausible but which is also such that if we set it aside, then either some or all of p_1, p_2, or p_3 would be less than plausible. This is perfectly conceivable. It is a prima facie challenge to the existence of any Goldilocks-esque middle ground.

This basic problem for (E) is pretty robust; it is hard to see how one could modify or interpret (E) in a way that will strike the right balance between setting aside too much evidence and too little evidence in all cases (of paradox). But perhaps I am simply lacking imagination. Consider the following proposed defense of (E).

> We can understand any paradox as a *reductio ad absurdam*. For example, consider a paradoxical argument—choose your favorite—of

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

> the form p, if p then q, q. We can understand this as a *reductio*, including the negation of the conclusion in the premises as follows: p, if p then q, not-q, therefore $\perp$. Now proceed step by step, premise by premise, as follows. Begin by considering the first premise, if p then q, on its own, only setting aside the available evidence for the remaining premises, p and not-q. Now consider the second premise, p, on its own, only setting aside the available evidence for the first and third premises, if p then q and not-q. Now, consider the third premise, not-q, on its own, only setting aside the evidence we have for, if p then q and p. If each of these premises is sufficiently probable given the evidence, then the argument is, by (E), a paradox. This step-by-step, premise-by-premise strategy allows us to set aside the right amount of evidence—neither too much nor too little—and so to salvage (E). Problem solved.

This imaginative defense of (E) does not succeed either. It still falls foul of the same basic problem (i.e., striking the right balance between setting aside too much evidence and too little). To see this, consider our schematic example above: a paradox—arranged as a *reductio*—of the form p, if p then q, not-q therefore $\perp$. Now presumably there could be some values for these premises such that there is some piece of evidence, e, that is evidence for both the first premise, if p then q, and the second premise p, and without which neither of these two premises would be very plausible. Suppose this is in fact the case. Now suppose we follow the proposed defense of (E) above. We begin by assessing the first premise, *if p then q*, setting aside the evidence for the second and third premises. *Ex hypothesi*, setting aside the evidence for the second premise entails setting aside e. So we must set aside e when assessing if p then q. But—also *ex hypothesi*—if p then q is less than very plausible if assessed setting aside e. It follows that our argument—p, if p then q, not-q therefore $\perp$—is not a paradox. But it is a paradox. Thus, the proposed defense of (E) fails. In order to counter this it would be necessary to argue that there can be no paradox in which there is a piece of evidence that is common ground between the premises and the negation of the conclusion. But it is unclear how or why one would argue in this way.

This rebuttal has been sketched at a high level of abstraction. But there are actual paradoxes that illustrate the point: paradoxes in which some claim, or piece of evidence, is an indispensable part of the case for more than one premise. Consider, for example, one version of Parfit's famous 'mere addition paradox' (Parfit 1984). We begin with a small, perfectly equal, population of people all with high levels of welfare. Call this A. At the first stage, we increase their welfare levels and add some more people, also with a positive welfare level, though lower than that of those at A. Call the resulting population A+. We can be pretty confident that A+ is not worse than A; after all, there is more of whatever makes life worth living at A+ than at A, and nobody who exists at both A and A+ is any worse off at the latter. At the second stage, we now modify the welfare levels of everyone at A+ so that the resulting population, call it B, has a higher total welfare, average welfare, and equality level than A+, but a lower average welfare level. We can be pretty confident that B is no worse than A+; after all, it has a

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

higher total welfare, average welfare, and equality level. By the transitivity of 'not worse than' we can be pretty confident that *B* is not worse than *A*. Yet, we repeat this process, and we end up with the implausible result that there is some population, *Z*, in which everyone lives a life that is barely worth living, but which is better than *A*. Hence the paradox. In this paradox, *both* our confidence that *A*+ is not worse than *A* *and* our confidence that *B* is not worse than *A*+ rely in part on the common claim that one population having more of whatever makes life worth living than another is a *pro tanto* reason for thinking the former better than the latter. The abstract point made earlier—that there could be paradoxes in which some claim or piece of evidence is an indispensable part of the case for more than one premise—is here illustrated in concrete terms.

4. Functionalist Accounts

The first accounts we considered, (A) and (B), were either too subjective or too objective. While (C) looks like a happy middle ground, it is problematic too. Jumble arguments threaten to show that it is too permissive, and the problem of paradoxical involvement pushes in the opposite direction, showing that it is too restrictive. Modifications of (C),namely, (D) and (E),do not do enough to resolve this. Where should we go from here? Perhaps we should persist with modifications to (C), but when the epicycles become tortuous, it can sometimes be worth changing one's approach altogether. Accounts like (C), including (D) and (E), are what we might call standard or orthodox accounts of paradox. They tell us to understand paradoxes in terms of the epistemic properties of the propositions that comprise them and the epistemic relations between those properties. Perhaps this is the wrong way to think. Perhaps we should instead understand paradoxes in terms of what they do: a paradox is an argument that gets us—or should get us—to think in some way or to revise our thought in some way. I call these functionalist accounts.

To the best of my knowledge this distinction has not been discussed in the literature, nor have functionalist accounts been seriously pursued. This is surprising. It is natural to think of paradoxes as a species of *puzzle*; and for puzzles, a functionalist account arguably *is* the obvious place to look. What it is to be a puzzle is to be the kind of thing that makes us react in a certain way. A functionalist approach to paradoxes is a sensible place to look too. Perhaps this way of understanding paradoxes speaks to the worry with jumble arguments. Jumble arguments are not paradoxes precisely because they do not puzzle us—or at least not in the right way. There is nothing puzzling about a random set of 100 propositions that you have recently uttered being inconsistent and therefore entailing anything.

So what exactly, on a functionalist view, is a paradox? Let us start with something simple: a paradox is the kind of argument that puzzles someone. This is obviously too simple. It is too permissive. Imagine that an argument puzzles me because it is written in a peculiar way so that it causes an optical illusion; perhaps its drawn in an Escher-like fashion so that its premises climb up to a conclusion that appears to be below them. This is not a paradox. We need a

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

more restrictive account. The point is quite general. Lots of functions are too permissive. Consider, for example, Sorensen's account, offered briefly in the preface to his book: 'Paradoxes are questions that suspend us between too many good answers' (2003: xii).

This is functionalist; a paradox is something that does something (suspends us). But it is also clearly too permissive. Suppose you ask me: 'What would you like for dinner—steak and chips or some roast pork?' There is a sense in which this is a question with too many good answers. Either answer (steak or pork) is great, but I cannot have both. I may therefore be unable to make up my mind: suspended between too many good answers. Clearly though, this is not a paradox. We need something more restrictive.

One appealing kind of option is normative functionalism: paradoxes are arguments that should make us do certain things. Perhaps, for example, a paradox is an argument that requires someone to reject either a premise, the reasoning, or the negation of the conclusion. This account does not entail that Escher-like drawings of arguments are paradoxes (or that questions about mealtime are paradoxes), but it is still too permissive. Suppose an evil demon presents you with a seemingly sound argument and tells you that he will destroy the world unless you reject a premise, the reasoning, or the negation of the conclusion. Presumably you are now required to do just this. Again though, this is not a paradox. Perhaps, then, a paradox is an argument that requires one to reject either a premise or the negation of the conclusion on epistemic grounds. This is much better. It avoids the problem with the evil demon; the grounds on which the demon forced you to choose were practical not epistemic. Clearly though, this account of paradox is also pretty skeletal. The qualifier 'on epistemic grounds' is really just a placeholder for a more substantive account; we would need to know what the referenced 'epistemic grounds' actually are. The obvious option is to understand them in terms of consistency, the thought being that a paradox is an argument in which one is required to choose between the premises and the negation of the conclusion on pain of inconsistency. This is still too permissive. It overgenerates paradoxes. There can be arguments that are not paradoxes but which are nevertheless such that one is required on grounds of consistency to choose between the premises and the negation of the conclusion. Consider any old argument of the form: p, if p then q, q. I am required on grounds of consistency to choose between the premises or the negation of the conclusion; I cannot consistently hold both. But it does not follow that any modus ponens is a paradox. Whether it is depends on the contents of 'p' and 'q'. Unless they have the right properties, it will not be a paradox.

So what are 'the right properties'? The overwhelmingly obvious thing to say is that both the premises and the negation of the conclusion would have to be plausible. When an argument is like that, as some but not all are, there is a paradox. But now look what has happened. We have come full circle. In attempting to find the right function, we have ended up with an account that looks very much like (C) (or (D) or (E)). What has happened?

Take a step back. Functionalist accounts tell us what paradoxes are in terms of what they do. If such accounts are to stand a chance, they are going to have to be quite fine-grained with respect to what gets 'done'. They cannot just say 'paradoxes are the kinds of things that puzzle' or 'paradoxes are the kind of thing that require you to revise your beliefs'. These are far too coarse-grained and permissive; there are too many things that are not paradoxes that perform these functions. Now a worry emerges, a worry that the brief discussion above is intended to illustrate. In order to get the requisite fineness of grain, functionalist accounts are going to end up having to incorporate standard or orthodox accounts directly within them. The slide toward this was illustrated above. In trying to work out how exactly a paradox puzzles, we were led to the view that it puzzles in the distinctive way of—something like—requiring one to either give up plausible premises or accept an implausible conclusion entailed by those premises. But that is just building something very like the standard or orthodox account into the functionalist view. Consider another way of putting the point. To get a sufficiently fine-grained account the functionalist is going to have to tell us that a paradox is an argument in which one does or is required to do something, in light of something else, only if something else, setting aside something else. . . . In filling out the function that paradoxes perform with a sufficient fineness of grain to make the account capture all and only paradoxes, the functionalist is effectively going to have to fill in these 'somethings' using exactly the materials that the standard or orthodox accounts use.

Why is this any of this a *problem*? If we are looking to functionalism to resolve the problems with standard or orthodox accounts—as we are; recall that we are trying an alternative approach to escape the epicycles—it seems that we will be disappointed. We will be disappointed because the functionalist is going to face exactly the same problems that standard or orthodox accounts face. This is for the simple reason that functionalist accounts will *include* standard or orthodox accounts.

To think about this worries in more detail, let us work with an example. Suppose that one were to propose the following functionalist account:

> (F) A paradox is an argument that engenders puzzlement in virtue of its premises being very plausible, yet entailing a conclusion that is very implausible.

The worry is that this account will run into the very same problems that we have already seen standard or orthodox accounts—i.e., (C), (D) and (E)—run in to. Most obviously, (F) presupposes that paradoxes do have very plausible premises and very implausible conclusions. But the problem of paradoxical involvement challenges this. Not all paradoxes do have very plausible premises and very implausible conclusions. Perhaps, then, a functionalist could modify (F) in some way to get around this. She could add in a modification to the effect that the premises are *individually* very plausible (G), or that they are very plausible *setting to one side* the fact that they entail the conclusion (H). But these are just the modifications familiar from (D) and (E) as discussed earlier. And as we say earlier, they do not work. It seems that moving from standard or orthodox accounts to a functionalist account isn't helping. It is just landing us back in the same, familiar problems. This is a consequence of the fact that for the functionalist account to be sufficiently fine-grained to capture all and only paradoxes it must incorporate a standard or orthodox account.

This requires one interesting qualification. While functionalist accounts like (F) (G) and (H) may be just as vulnerable to the problem of paradoxical involvement as their standard or orthodox equivalents (C), (D) and (E), they are arguably at least somewhat more resilient to jumble arguments. The issue with jumble arguments—as noted at the beginning of this section—is precisely that they *do not* engender puzzle. There's nothing *puzzling* about a random set of 100 propositions that you have recently uttered being inconsistent and therefore entailing anything. Now (F) (and (G) and (H)) explicitly states that paradoxes are arguments that do puzzle. So it entails that jumble arguments are not paradoxes. Here the functionalist dimension of (F) is doing some useful work. The work is only *somewhat* useful though. It leaves the interesting and difficult work undone. (F) does not tell us *why* some arguments with plausible premises and implausible conclusions (or some suitable modification of this) puzzle and why others – like jumble arguments – do not. That is what we would really want to a fully satisfactory account of paradox to explain. In fact, the problem is even harder than this. How could it be that some arguments puzzle simply in virtue of having some property (i.e. plausible premises, implausible conclusion), but that other arguments that have the very same property do not puzzle? It is prima facie unclear how this could be. So it is unclear that (F)—or functionalist accounts more generally—really does make any progress with respect to jumble arguments. And it certainly does not make any progress with respect to the problem of paradoxical involvement.

5. A Puzzle About Paradoxes

We now have the beginnings of a puzzle. Standard or orthodox accounts of paradox allow us to find middle-ground between the subjective and the objective. But they are vulnerable to both jumble arguments and, I have claimed, the problem of paradoxical involvement. Functionalist accounts represent a promising alternative. But it turns out that they inherit the very same problems. So it is simply not obvious what paradoxes are. The problem worsens if we think that standard and functionalist accounts come close to exhausting the conceptual space. Standard accounts understand paradoxes in terms of their 'internal' properties. Functionalist accounts understand paradoxes in terms of their 'external' properties. What other options could there be? The result may not rise to the level of paradox itself but it is certainly something worth puzzling over.

CHRISTOPHER COWIE
UNIVERSITY OF DURHAM, UK
christopher.d.cowie@durham

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press

References

Clark, Michael. (2012) *Paradoxes: From A to Z*. London: Routledge.

Lycan, William. (2010) 'What, Exactly, Is a Paradox?' *Analysis,* 70, 615–22.

Parfit, Derek. (1984) *Reasons and Persons*. Oxford: Oxford University Press.

Priest, Graham. (2006). *In Contradiction*. (2nd ed.). Oxford: Oxford University Press.

Priest, Graham, J. C. Beall, and Bradley Armour-Garb, eds. (2004) *The Law of Non-Contradiction*. Oxford: Oxford University Press.

Quine, W. V. O. (1962) 'Paradox'. *Scientific American,* 206, 84–99.

Rescher, Nicholas. (2001) *Paradoxes: Their Roots, Range and Resolution*. Chicago: Open Court.

Sainsbury, Mark. (1988) *Paradoxes*. 1st ed. Cambridge: Cambridge University Press.

Sorensen, Roy. (2003) *A Brief History of the Paradox: Philosophy and the Labyrinths of the Mind*. New York: Oxford University Press.

Temkin, Larry. (2012) *Rethinking the Good*. New York: Oxford University Press.

https://doi.org/10.1017/apa.2021.48 Published online by Cambridge University Press