Acta Analytica (2023) 38:211–228
https://doi.org/10.1007/s12136-022-00515-2

Some Remarks on the Notion of Paradox

Sergi Oms[1]

Received: 21 June 2021 / Accepted: 18 March 2022 / Published online: 5 July 2022

Abstract
This paper argues that the traditional characterization of the notion of paradox — an apparently valid argument with apparently true premises and an apparently false conclusion — is too narrow; there are paradoxes that do not satisfy it. After discussing, and discarding, some alternatives, an outline of a new characterization of the notion of paradox is presented. A paradox is found to be an apparently valid argument such that, apparently, it does not present the kind of commitment to the conclusion that should be implied by an acceptance of the truth of the premises and the validity of the argument.

Keywords Paradox · Curry's paradox · Liar paradox · Sorites · Normativity of logic

Traditionally, an argument has been considered a paradox if, and only if:

(i) it is an apparently valid argument,
(ii) it has apparently true premises, and
(iii) it has an apparently false conclusion.

This view can be found, for instance, in Quine (1966, page 7), Beall & van Fraassen (2003, page 119), Priest (2006, page 9), Cave (2009, page 3), Sainsbury

✉ Sergi Oms
sergi.oms@ub.edu

1 Facultat de Filosofia, University of Barcelona, Logos Group, BIAP, C/ Montalegre, 6, 08001, Barcelona, Spain

(2009, page 1) and Armour-Garb (2017, page 3) among many others.[1] This paper argues that the traditional characterization of the notion of paradox is too narrow — to wit, the conditions listed above are not necessary for being a paradox — and introduces an alternative proposal.

In the first section, I argue that the traditional characterization of the notion of paradox does not apply to certain kinds of arguments that are problematic in the same way as arguments that do fit the traditional characterization. Hence, the traditional characterization fails to capture whatever these two kinds of arguments have in common. Sections 2 and 3 discuss two alternative proposals to the characterization of the notion of paradox. In Section 4, I present a more general characterization of the notion of paradox that includes both the arguments that satisfy the traditional characterization and those arguments which, even if they do not satisfy it, are problematic in the same way. Finally, Section 5 contains some concluding remarks.[2]

1 The Traditional Definition

Let us state again the traditional definition of the notion of paradox, henceforth ‘Definition 1’:

Definition 1 A paradox is an apparently valid argument with apparently true premises and an apparently false conclusion.

It is important to note that the sense in which ‘apparently valid’ (‘true’, ‘false’) is used in this definition is quite strong—although not as strong as declaring it valid (true, false), of course. To wit, a paradox is not an apparently valid argument in the sense that it merely *seems* valid, but in the sense that, declaring it invalid, implies giving up strong intuitions about logic. And the same occurs with the other conditions of the traditional characterization. Thus, the premises of a paradox are apparently true in the sense that their being not true would violate some of our core intuitions with respect to some of the concepts—either explicitly or implicitly—involved in them

[1] Of course, some variations occur: for example, sometimes absurdity (Quine), untruth (Beall and van Fraassen), unacceptability (Sainsbury), high counter-intuitiveness (Cave) and even being a dialetheia (Priest) are mentioned. For ease of explanation, I will simplify and I will take all these notions to imply apparent falsity. On other occasions, the definition focuses on the conclusion rather than on the argument leading to it (e.g., Quine or Sainsbury). Again, I will simplify and take the above definition to encompass all the aforementioned variations. (Nothing essential hinges on this decision.) Other authors have proposed to charaterizing paradoxes as sets of mutually incompatible, or inconsistent, truth-bearers (see Schiffer 2003, p. 68, Lycan 2010, p. 618 or Sorensen 2003, pp. 6, 364). As I argue in footnote 10, the considerations presented in this paper also apply to these characterizations made in terms of sets. See Sorensen (2003, p. 3-10) or Armour-Garb (2017, p. 3) for finer taxonomies of the characterizations of the notion of paradox available in the literature.

[2] While this paper was under review, Zardini (2021) has presented a new approach to the notion of paradox according to which a paradox is ‘a situation where, apparently, even if the conclusion failed to hold, all the elements of the putative proof of the conclusion would still be available’ (p. 505). I defer any further discussion of Zardini’s proposal to future research.

(and the same can be said about the apparent falsehood of the conclusion).[3] Accordingly, the sense of ‘paradox’ that I am trying to help elucidate in this paper is the technical or semi-technical sense that is usually used in the literature in philosophy of logic and language (what Quine 1966 calls ‘antinomy’). In this sense, then, a paradox points to a tension in the basic intuitions governing one or more of our concepts. That is why solving a paradox must involve giving up some of these core intuitions, all of which are, typically, equally cherished, and as a result, there is no agreement whatsoever as to which of them has to be abandoned. So the apparent validity of the argument, the apparent truth of the premises, and the apparent falsity of the conclusion are *forced* by the tension that arises from the concepts involved in the paradox; that is, we would be willing to state that the argument is valid, that the premises are true and that the conclusion is false but since this is, *in principle*, impossible, we are forced to declare these properties as *apparent*.[4]

I want to show next that Definition 1 is too narrow. I will offer two arguments that I claim can reasonably be considered paradoxes but that do not satisfy Definition 1. The first counterexample to Definition 1 I want to consider is Curry’s paradox.[5] It can be presented concisely as follows.

Suppose that T is the truth predicate and that it obeys the so called ‘T-schema’, according to which, for any sentence ϕ, $T\ulcorner\phi\urcorner \leftrightarrow \phi$.[6] Suppose, furthermore, that we have a Curry sentence γ, which is a sentence that asserts that if itself is true then snow is white:

$$(\gamma) T\ulcorner\gamma\urcorner \rightarrow \text{snow is white}$$

Suppose, now, that $T\ulcorner\gamma\urcorner$. Then, under this supposition, what γ says is the case; to wit, if $T\ulcorner\gamma\urcorner$ then snow is white. We can apply next *modus ponens* and conclude, under the assumption that $T\ulcorner\gamma\urcorner$, that snow is white. Since we have concluded that snow is white under the supposition that $T\ulcorner\gamma\urcorner$, we can now conclude that *if* $T\ulcorner\gamma\urcorner$, *then* snow is white. So we just proved γ itself and, hence, $T\ulcorner\gamma\urcorner$. At this point, then, we have

3For more discussion on the meaning of ‘apparently’ in paradox related discussions see Priest (2003, pp. 277–78) and Oms and Zardini (2021, fn. 2).

4This suggests, as far as I can see, one of the main differences between a *paradox* and a *fallacy*; only in the latter case we take for granted that there is an error in the argument, while in the former it is left open whether or not to accept the argument and its conclusion. Of course, in some cases, *solving* a paradox might involve declaring it a fallacy, but still, even in these cases, it has to be argued that the paradox is a fallacy, it cannot be merely presupposed—the idea that fallacies are some kind of error in the argumentation can be found, for example, in Powers (1995, p. 308), Walton (1989, p.19), Walton (1995, p. 255) and Ikuenobe (2014, p. 210), among many others. Notice also that, given that a paradox points to a tension in the basic intuitions governing some of our concepts, if these concepts evolve and there is a change in some of their underlying intuitions, the paradox might disappear. So, as I am conceiving it here, the notion of paradox is relative to our understanding of the relevant concepts, which might change over time. Although I will not pursue this issue here, I think this might be the case, for example, with the so-called Galileo’s Paradox and, *perhaps*, also Zeno’s Dichotomy, which would no longer be paradoxes (in the aforementioned sense). (Of course, the notion of paradox is a vague one.) Thanks to Jose Díez, Manuel García-Carpintero and an anonymous referee for important insights on this issue.

5To the best of my knowledge, the idea of using Curry’s paradox as a counterexample to Definition 1 was first proposed by López de Sa and Zardini (2007).

6I will use the brackets (‘$\ulcorner$’ and ‘$\urcorner$’) in order to indicate some device of canonical name formation for sentences; thus, ‘$\ulcorner\phi\urcorner$’ is just a name for the sentence ϕ and ‘T$\ulcorner\phi\urcorner$’ ascribes truth to the sentence ϕ.

the following two claims: first, that if $T\ulcorner\gamma\urcorner$, then snow is white; and, second, that $T\ulcorner\gamma\urcorner$. From these, applying *modus ponens* again, we conclude that snow is white.

This is, roughly put, how Curry's Paradox allows us to conclude the consequent of γ; thus, in this particular case, it allows us to conclude that snow is white. Notice now that according to Definition 1, the argument provided by Curry's reasoning with the use of γ is not a paradox, for the conclusion obtained is not false. I claim that this argument, though, is a paradox. Let me elaborate on that.

In general, Curry's Paradox can be presented with a variable ϕ ranging over sentences in the position of 'snow is white' in γ; then, it would not have been a proper argument, but a schema whose instances would have been arguments. Independently of the interpretation of the sentence variable in Curry's sentence, all instances of the schema would have been paradoxes. But, since ϕ could have been a true sentence, we conclude that we will have paradoxes, like the one mentioned above, with true conclusions. Hence, paradoxes like these are counterexamples to Definition 1.

At this point, there is a natural rejoinder that a proponent of Definition 1 could give in response to Curry's paradox. A defender of Definition 1 could say that when we use a certain sentence ϕ to formulate Curry's paradox the unacceptable conclusion we achieve is not ϕ but that ϕ logically follows from the premises in the Curry's reasoning. We would still have, then, an apparently valid argument with apparently true premises and an apparently false conclusion (namely, the claim that ϕ logically follows from the premises in Curry's argumentation). This reply might work with true sentences like 'snow is white'; that is, if we run Curry's paradox with 'snow is white' in the Curry sentence and we conclude 'snow is white', we can read the paradox as concluding that 'snow is white' logically follows from the premises in Curry's reasoning.[7] Since this last claim is apparently false, Definition 1 would be vindicated. But notice that, even granting that understanding of Curry's paradox, we can also build a Curry paradox using a logically valid sentence or a true arithmetical sentence, say, ψ. Then even if we understand Curry's paradox as having as its conclusion that ψ logically follows from its premises, this conclusion will no longer be apparently false, but plainly true; because, in the case of ψ being a valid sentence, it will logically follow from the premises in Curry's reasoning (in fact, it will follow from any premises) and in the case where ψ is a true arithmetic sentence, since arithmetic is present in the premises to prove the Diagonal Lemma (at least in some of the ways for formulating Curry's paradox), ψ will logically follow from them.[8]

In conclusion, even assuming that some of Curry's paradoxes can be understood in a way such that they are no longer counterexamples to Definition 1, we can still, with the use of logically valid or true arithmetical sentences, devise other Curry's paradoxes that are.

[7]But note that the paradoxical result seems to have been achieved before that; to wit, we already are perplexed when we obtain ϕ as the conclusion of our argument, which suggests that the meta-claim according to which ϕ should not follow from the premises is not essential to Curry's paradoxicality.

[8]It could still be said that what makes Curry's argument a paradox is the fact that it trivializes the logic, since it allows us to conclude anything (the same would occur for any paradox that has a contradiction as its conclusion, if we accept explosion). But trivialization of the logic cannot be a good way of defining what a paradox is, because there are paradoxes, like the Sorites, that do not trivialize the logic.

Let us present, next, another counterexample to Definition 1 which also shows that it does not provide the necessary conditions for being a paradox. Consider, for example, the following argument, where Alice is 120 cm tall (a clear case of not being tall):

1. Alice is tall,
2. if Alice is tall, so is someone who is 1 cm shorter,
3. someone 1 cm shorter than Alice is tall,
4. if someone who is 1 cm shorter than Alice is tall, so is someone who is 2 cm shorter than Alice,

   ⋮

151. if someone who is 99 cm shorter than Alice is tall, so is someone who is 100 cm shorter than Alice,
152. hence, someone who is 100 cm shorter than Alice is tall.

According to Definition 1, this is not a paradox; for it would only be a paradox, a Sorites paradox,[9] in situations where Alice was, say, 200 cm, as, then, the premises would be apparently true and the conclusion (that someone who is 100 cm is tall) would be apparently false. My point is that, even when Alice is not tall, the argument is still a paradox (see the discussion below on the paradox/pathodox distinction). Hence, once again, we conclude that the conditions stated in Definition 1 are not necessary for being a paradox. Notice that counterexamples of the soritical kind, such as Argument 1, are the strongest ones that can be raised against Definition 1 and, accordingly, they are perhaps also the least plausible ones. As a matter of fact, though, I think other weaker (and more plausible) soritical counterexamples can be presented. Let us see how.

As a number of authors note (see, especially, Barnes 1982, page 30), in order to construct a Sorites paradox with a given vague predicate P, it is sufficient to have an ordered series of objects $a_1, a_2, \ldots a_n$ such that Pa_1, $\neg Pa_n$ and such that all adjacent objects in the series must be related by the tolerance relation (see Wright 1975, page 333); that is, for each $0 \leq i < n$, a_i and a_{i+1} must be indiscriminable with respect to the application of P. What I am claiming is that although these conditions are sufficient for having a paradox, they are not necessary, as we can still have a paradox when $\neg Pa_1$ and $\neg Pa_n$ (and, indeed, when Pa_1 and Pa_n). We can even construct a series of objects $a_1, a_2, \ldots a_n$ respecting the tolerance relation as above such that Pa_1, Pa_n and $\neg Pa_m$, where $1 < m < n$. In this case, the corresponding Sorites argument would begin with the true claim that a_1 is P; it would, then, argue the false claim that a_m is P; and, it would end with the true conclusion that a_n is P. Hence, this argument would not fit the traditional characterization of a paradox, for

[9] For a comprehensive introduction to the *Sorites* paradox, see Oms and Zardini (2019b).

it would have a true conclusion, in spite of the fact that its paradoxical character is hard to deny (cf. Oms & Zardini 2019a, fn. 14).[10]

At this point, the natural rejoinder in defense of Definition 1 is that the arguments just presented (like the Curry argument, where Curry's sentence is build using 'snow is white', and the Sorites argument where Alice is not tall) are not paradoxes—so that they are not valid counterexamples to Definition 1—, but something different, albeit that they are closely related to paradoxes. Let us call them, say, 'pathodoxes' (from *pathos* and *doxa*; arguments leading to some kind of ill-formed opinion) and let us call the arguments that fit the traditional characterization 'traditional paradoxes'. The complaintant might go on by saying that the traditional characterization was not intended to characterize pathodoxes, but only the traditional paradoxes. As far as I can see, though, the discussion at this point is just a semantic one; if we really have to differentiate between these two notions—traditional paradoxes and pathodoxes—, then in this paper I seek a characterization of the notion that includes both, which I will call 'paradox'. I think this project makes sense for the following reasons.

The diagnosis of the problems that a traditional paradox and an analogous pathodox pose seems to be the same, as does the solution. To wit, if somebody claims that a traditional Sorites paradox should be solved in a different way from a Sorites pathodox, we would expect some explanations as to why this should be the case, for the conceptual tensions to which both arguments give rise are the same. This means that the phenomena underlying both kinds of arguments are the same and that research involving both traditional paradoxes and pathodoxes must help enlighten the notions involved in them in the same way; hence, the importance of capturing the notion that encompasses both traditional paradoxes and pathodoxes.

To be clear, I am not claiming that pathodoxes are paradoxes because they instantiate some schemata or patterns of reasoning which are instantiated respectively by traditional versions of the paradoxes in question (as a matter of fact, I explicitly deny this claim in the next section). What I am saying is that a pathodox and an analogous traditional paradox are both problematic; there is something wrong with both of

[10] As mentioned in footnote 1, there is another interesting proposal for a definition of the notion of paradox that should not be overlooked. Lycan (2010) proposes the following characterization of the notion of paradox:

Definition 1' A paradox is an inconsistent set of [sentences], each of which is very plausible. (Lycan, 2010, 618)

According to Lycan, a paradox is typically obtained by putting together the premises and the negation of the conclusion in a set, so that the result is an inconsistent set (because of the paradoxical argument) with plausible sentences as elements (Lycan, 2010, p. 617). Lycan's definition, however, falls prey to the same problems as Definition 1, since Lycan is supposing that the premises are acceptable and the conclusion is not (so that its negation is). Then, a Curry paradox that uses 'snow is white' in the Curry sentence might look something close to the following (where Γ is the set of premises):

$$\{\Gamma,\ \text{snow is not white}\},$$

which clearly does not count as a paradox according to Lycan's proposal because, although it is an inconsistent set, not all of its members are plausible. So, as in the case of Definition 1, Lycan's definition does not offer necessary conditions for being a paradox. To my mind, the same considerations apply to all the other approaches to paradoxes that define them in terms of sets of truth-bearers as introduced in footnote 1.

them. Moreover, the source of its problematic character seems to be the same in both cases. And, hence, it is natural to try to capture this problematic character in general for both pathodoxes and traditional paradoxes. In this paper, I want to explore some ideas regarding this general notion of paradox.

Albeit, perhaps, less persuasive, I think we can also appeal to the phenomenology of paradoxes; when we are faced with a paradox we have a characteristic feeling that there is something wrong, that is, a feeling which constitutes what it is like to be faced with a paradox, although we find it very hard to say exactly what that feeling is. This feeling is the same, for instance, with respect to any Curry case, regardless of whether the conclusion is acceptable or not. (It's just that when the conclusion is not acceptable, this feeling might be more pressing.)[11]

2 The Logical Form

One possible and, at first sight, natural alternative characterization of the notion of paradox could be stated along the following lines:

Definition 2 A paradox is an apparently valid argument whose *logical form* can be used to derive an apparently false conclusion from apparently true premises.

According to this definition, the pathological Sorites of Section 1 is a paradox even when Alice is tall, because an argument with the same logical form could be used to get a false conclusion from true premises. And the same with Curry's paradox using a true sentence to build Curry's sentence.

Definition 2, though, is too broad, for compare the following two arguments:

Argument 1

1. 2 is a natural number,
2. if a number is a natural number, so it is its successor,
3. hence, 20564 is a natural number.

Argument 2

1. 2 grains of sand do not form a heap,
2. if n grains of sand do not form a heap, neither do $n + 1$ grains of sand,
3. hence, 20564 grains of sand do not form a heap.

which have the same logical form.

[11] Although I will not pursue this point here, it is worth noticing that, given the Sorites pathodox introduced in Section 1, some soritical pathodoxes are such that their main inductive premise is true, not because of the vagueness of the predicate and its tolerance, but because of the fact that all the objects in the soritical series are clear countercases of the vague predicate used in the paradox, so that the ascription of such a predicate to them is always false and, hence, the conditionals with this ascription as antecedent are always true. This means that solutions that point to the non-truth of the inductive premise in the Sorites (like, for example, Supervaluationism and Epistemicism) cannot solve this kind of Sorites pathodoxes. Hence, these considerations would constitute an advantage for accounts of the Sorites paradox that point to some structural problem in the argument itself (like, for example, Zardini 2008 or Slaney 2010).

Now, according to Definition 2, since Argument 2 allows us to infer an apparently false conclusion from apparently true premises and since Argument 1 and 2 share the same logical form, we should claim that Argument 1 is a paradox; but, this is not the case. Clearly, the paradoxicality of Argument 2 does not depend solely on its logical form, but also on certain properties of the vague predicate ‘heap’.

At this point, an advocate of Definition 2 might try to argue that, in this case, there is a broader notion of form involved, which takes into account the vagueness of ‘heap’. Then, she would argue, the stipulation that the predicate used in the argument is vague should be understood as an intrinsic part of its form, in which case Arguments 1 and 2 would no longer have the same form. There are at least two problems with this line of thought. First, such an account of what a paradox is would presuppose a conception of vagueness that should be prior to its susceptibility to soritical arguments and, second, we would need to characterize the notion of form in question, which seems far from simple.[12]

3 A First Attempt

Another way out of this situation has been proposed by López de Sa and Zardini (2007):

Definition 3 What really seems to be of the essence [of a paradox] is that, despite the apparent validity of the argument, the premises do not appear rationally to support the conclusion. (López de Sa & Zardini 2007, page 67)[13]

This definition, though, is unclear in a way that could result in it, once again, being too broad. Consider the following argument:[14]

Zebra Argument

1. This is a zebra,
2. if this is a zebra, then it is not a cleverly disguised mule,
3. hence, this is not a cleverly disguised mule.

In some reasonable sense of ‘not rationally supporting’, the Zebra Argument just introduced is a valid argument (it is an instance of *modus ponens*) such that the premises do not appear rationally to support the conclusion. This argument is a prototypical case of an argument that begs the question. It seems that the Zebra Argument

[12] Thanks to an anonymous referee for prompting this clarification.

[13] A similar proposal has been defended by Cook (2013, p. 11), according to whom, a Curry’s paradox with a true sentence such as the one used above has a conclusion that is ‘merely inappropriate’ and, according to his definition of what a paradox is, this is enough to be paradoxical (p. 10). Hence, claims Cook, in some cases paradoxes are arguments that ‘lead to conclusions that might be true, but which, in some very real sense, should not follow from the premises in question’ (p. 11). To my mind, López de Sa and Zardini present a way to spell out why the conclusion should not follow from the premises (why it is inappropriate) and, hence, I take their proposal to offer greater clarification than Cook’s. That is also why in this paper I opt to focus on the former.

[14] Thanks to Manuel Pérez Otero for suggesting a counterexample of this kind.

begs the question because someone who does not accept the conclusion 3 will deny the evidence that supports 1—for instance, someone who thinks, precisely, that what seems a zebra is a disguised mule.[15]

What I wish to stress here is that if we understand the notion of *not rationally supporting* as something on the lines of *not giving the right kind of reason*—which is usually taken to be one of the features of begging the question arguments; see Sinnott-Armstrong (2012, page 179)—, something that can be typically tested in terms of *not succeeding dialectically*, then Definition 3 is too broad. For arguments like that of the Zebra Argument, which are not paradoxical, will count then as cases in which the premises do not rationally support the conclusion. Even more, plain circular arguments are also arguments such that the premises do not rationally support the conclusion *in the sense just stated*:

Circular Argument

1. snow is white
2. hence, snow is white

which means that this kind of argument would count as a paradox, too, according to this understanding of Definition 3 (with the aforementioned sense of 'not rationally supporting').

A proponent of Definition 3 could reply that the Zebra Argument and the Circular Argument are paradoxes, in particular, they are some kind of pathodox. It should be noticed, however, that circular arguments like the one above do not necessarily point to any tension in the concepts involved (furthermore, they do not seem to share the phenomenology of paradoxes, that is, when faced with them we do not feel the discomfort we feel when we are faced with a paradox). Let me elaborate on this in order to see why circular arguments are not, in general, pathodoxes. As we shall see, this distinction will help determine the definition of the notion of paradox I defend in this paper.

Take a pathodoxical Sorites like the ones presented in Section 1, and take, also, the Circular Argument. Notice that both are apparently valid arguments in the sense discussed before; to wit, declaring them invalid would require giving up core intuitions of the notion of logical validity. On the other hand—at least if we suppose that logic is normative; more on this below—, in each of them, in virtue of their validity, if I believe its premises I ought to believe its conclusion.[16] This can be stated in terms of commitment; in both the pathodoxical Sorites and the Circular Argument, if a subject believes the premises and believes the argument to be valid then she is committed, in virtue of the fact that she believes the premises and the fact that she

[15]The *loci classici* for the discussion on the notion of begging the question are Jackson (1984) and Walton (1991).

[16]For readability, I am being a bit loose with the scope of 'ought'; properly, I should have stated that, in each of the arguments, in virtue of their validity, I ought to see to it that if the premises are the case so is the conclusion.

believes the argument to be valid, to believe the conclusion.[17] The crucial difference between the pathodoxical Sorites and the Circular Argument is that when faced with the former I do not want to have to believe the conclusion in virtue of the fact that I believe the argument to be valid and the fact that I believe the premises; this commitment makes me uncomfortable. In contrast, in the latter case, I am willing to believe the conclusion in virtue of the fact that I believe the argument to be valid and the fact that I believe the premises; this commitment I embrace willingly. The same occurs with arguments that beg the question, such as the Zebra Argument; if I believe the premises and I accept the argument as valid, I willingly embrace the commitment to the conclusion. To my mind, the main characteristic of paradoxes and the reason of the phenomenology we associate with them is strongly related to the fact that we are not willing to accept the commitment that stems from them; this is what explains our discomfort when we are faced with a paradox. That is why we should not consider arguments like the Circular Argument or the Zebra Argument as paradoxes.

4 The Notion of Paradox

Nevertheless, the characterization provided by López de Sa and Zardini (2007) seems to follow the right track. We may try to refine it by making more precise about what is meant when it is said that in paradoxes the premises do not rationally support the conclusion.

We have seen that what differentiates traditional paradoxes and pathodoxes (to wit, paradoxes) on the one hand from question-begging and circular arguments on the other, is the fact that we are only willing to accept the commitment that follows from the apparent validity of the argument in the latter case. As a matter of fact, we can state something stronger. Consider, for instance, a pathodoxical Sorites with a false premise and a true conclusion (for example, like the one introduced at page 6, in which the true Pa_n is concluded through the false Pa_m). We have seen that, in this case, since it is still a paradox, we are not willing to accept the commitment that stems from it.

Suppose now that a given subject S who accepts the validity of the argument and the truth of the premises does not accept the truth of the conclusion. Suppose, also,

[17] As it stands, this claim presupposes an understanding of the normativity of logic which is too strong, for it implies, at the very least, that we are committed to believing infinitely many things. For simplicity, though, I will stick, especially in the formulation of principles C and (*) below, to the claim that if an agent believes the premises of an argument and she also believes the argument to be valid, then she ought to believe its conclusion. This means, first, that I am presupposing some idealizations on the cognitive resources available to the agent and, second, that the conclusion is such that it is being considered by the agent (so that the commitment does not apply to all the sentences that follow from the premises, just to those that are being considered by the agent). I hope then to avoid the commitment to believing infinitely many things. To my mind, as far as the characterization of the notion of paradox offered in this paper concerns, nothing essential hinges on these simplifications. How to formulate principles that capture the normative dimension of logic exactly is a highly controversial matter. (For some discussion see, for example, Harman 1984; Broome 1999, 2013; Field 2009 and Steinberger 2020) (and see the discussion in the Final Remarks below).

that nothing is known about the conclusion, in the sense that there are no any other arguments neither in favor nor against it, so that the subject is epistemically neutral with respect to it (let us call these conditions '*conditions of epistemic neutrality*'). If we are not willing to accept the commitment that the pathodox generates, then we will not be willing to accuse *S* of having done anything wrong.

Consider now the following principle:

C If there is a commitment to accept the conclusion of an argument by virtue of its validity and the acceptance of the truth of its premises, then (under conditions of epistemic neutrality) a subject who accepts the argument to be valid, the truth of the premises and does not accept the truth of the conclusion is doing something wrong.

Then, since, as we just saw, when faced with a pathodoxical Sorites we believe that a subject who accepts the argument to be valid, the truth of the premises, and does not accept the truth of the conclusion is not doing anything wrong, C implies that there is no commitment implied by the pathodoxical Sorites. This can be generalized to any paradoxical argument and, hence, any paradoxical argument is such that when we reflect on the notion of commitment and on how a subject should behave when faced with a paradox we realize that, apparently, there is no commitment to accept the conclusion in virtue of the acceptance of its validity and the acceptance of its premises (under conditions of epistemic neutrality). So, it is not only that we are not willing to accept the commitment that stems from a paradox (as we just saw), but that, apparently, there is no commitment at all![18]

Compare this situation with our discussion of traditional paradoxes and Definition 1 in Section 1. We said that, when faced with a traditional paradox we see that, although the argument is apparently valid, its premises are apparently true and its conclusion is apparently false (in the sense that denying any of these claims would involve giving up some core intuitions of either validity or some of the key notions in the argument). We saw that these must be apparent, because they are jointly impossible. At the same time, it could be claimed, since having true premises and a false conclusion is a sufficient condition for being invalid, a traditional paradox is an argument that is apparently valid and apparently invalid.

We are now in a similar situation. Traditional paradoxes and pathodoxes (to wit, paradoxes) are apparently valid arguments—and, hence, apparently, they make us commit, under conditions of epistemic neutrality, to the conclusion when we accept the validity of the argument and the truth of the premises—and, at the same time, apparently, there is no commitment at all (given the argument above involving principle C). Again, these must be apparent for they are arguably jointly impossible (see the Final Remarks though). Moreover, if generating commitment to the conclusion is

[18]To my mind, that is why, at least *in principle*, we cannot use a paradox as a *reductio* argument; in order to do so we need to dispel the apparent lack of commitment. The main reason behind what I called the phenomenology of paradoxes (Section 1) seems to be rooted in this fact: we cannot explain why, while believing the premises and accepting the validity of the argument, a subject *S* remains, *apparently*, uncommitted to the conclusion.

a necessary feature of any valid argument, then a paradox is, as before, an argument that is apparently valid and apparently invalid.

Let us try to spell this out more clearly. The idea is that for any given set of premises Γ that imply a sentence δ in a certain argument A and any subject S the following is the case:

(*) if S believes Γ and believes that A is valid, then S is committed (under conditions of epistemic neutrality) to believe δ.

We can then introduce the following definition of the notion of paradox:

Definition 4 A paradox is an *apparently* valid argument such that, *apparently*, (*) fails; that is, *apparently*, someone can believe the premises and believe that the argument is valid while not being committed (under conditions of epistemic neutrality) to the conclusion.

When faced with a paradox, we are not committed, apparently, to believe its conclusion even when, under conditions of epistemic neutrality, we believe that the argument is valid and that the premises are true.[19]

Therefore, the idea underpinning Definition 4 is that when faced with a paradox there are two strong, confronting appearances that make us reconsider some of our basic intuitions of some of the concepts somehow or other involved in the paradox. On the one hand, the rules that constitute our logic lead us to consider the paradox as a valid argument. But on the other, when we reflect on the commitments that follow from our acceptance of the premises and our acceptance of the validity of the argument, we realize that, apparently, there is no commitment at all.[20] And the appearances of validity and invalidity, as in the case of Definition 1 (see the discussion at the beginning of Section 1), are *forced* upon us.

Definition 4, thus, can be seen as a generalization of Definition 1. In the latter case, the conclusion of the paradox was something unacceptable, typically, a contradiction. Similarly, in the former case, there also is something apparently unacceptable, namely, the fact that a certain apparently valid argument apparently does not satisfy (*). As a matter of fact, if satisfying (*) is taken as a necessary condition for being a valid argument, we are then faced, as we saw, with an uncomfortable situation: the argument is apparently valid and apparently invalid.

[19] Cf. Oms (2020) for some of the previous ideas that led to Definition 4.

[20] To be clear, according to this approach, a hallmark of paradoxicality is this apparent lack of commitment. For instance, consider the case of a pathodoxical Sorites with a false premise and a false conclusion. In such a case, if I (falsely) believe the premises and I believe that the argument is valid, I am not, apparently, committed (under conditions of epistemic neutrality) to the conclusion, although, at the same time, it seems that I should be committed to the (false) conclusion (which would be the case if the argument was not a paradox). Thanks to an anonymous referee for prompting this clarification.

5 Final Remarks

Definition 4 can be seen as a generalization of Definition 1 for yet another reason. It is usually agreed that solving a traditional paradox involves one of the following options: (i) denying the validity of the argument (and explaining why it seems valid); (ii) denying the truth of some of the premises (and explaining why they seem true); or (iii) denying the falsity of the conclusion (and explaining why it seems false). As we will see, according to Definition 4, solving a paradox involves these very same options.

In order to see this, let us try to capture the essentials of the normative nature of valid arguments. In order to do so I will use what MacFarlane (2004) calls 'bridge principles', principles that try to connect logical facts with norms of reasoning. MacFarlane proposes an exhaustive list of such principles, but, for our purposes here it will be enough to use (a variation of) one of these bridge principles that has been endorsed by some authors for capturing the (alleged) normative nature of logic (it is adapted from Broome 1999. See also Steinberger 2020 and MacFarlane 2004). Suppose that Δ is an argument with the members of the set Γ as premises and δ as the conclusion. Then,

$$\mathrm{O}(\mathrm{B}\Gamma \rightsquigarrow \mathrm{B}\delta),$$

is taken to mean that you ought to see that if it is the case that you believe the premises in Γ then it is the case that you believe the conclusion, *under conditions of epistemic neutrality*.

According to Definition 4, then, a paradox is an apparently valid argument with premises Γ and conclusion δ, such that, apparently,

$$\neg\mathrm{O}(\mathrm{B}\Gamma \rightsquigarrow \mathrm{B}\delta)$$

We can show now what is needed to solve a paradox according to Definition 4. First, in what we will call a *type 1* solution, we can show that the argument is not valid. In this case, it would be immediately explained why believing the premises does not commit you, even under conditions of epistemic neutrality, to believe the conclusion. Ideally, we should be able to explain why, pace the fact that the argument is not valid, it is apparently valid, so why we should abandon the intuitions about validity that are involved in its being apparently valid and why they are so compelling.

Alternatively, in a *type 2* solution, we can defend the validity of the argument. In this case, since the argument is valid, believing the premises commits you, under conditions of epistemic neutrality, to believe the conclusion; or, more succinctly,

$$\mathrm{O}(\mathrm{B}\Gamma \rightsquigarrow \mathrm{B}\delta)$$

and hence, the failure of commitment to the conclusion must be just apparent. This appearance would be prompted, according to the proponents of a type 2 solution, by the fact that we are offered a case of an argument in which the premises are apparently true and the conclusion is apparently false. The situation can be described as follows. From the fact that, apparently,

1. $\mathrm{OB}\Gamma$

and, apparently,

2. $O\neg B\delta$

you conclude that, apparently,

3. $\neg O(B\Gamma \rightsquigarrow B\delta)$[21]

What a proponent of a type 2 solution would say is that it is this inference that explains why we think that a paradox is an apparently valid argument that, apparently, does not present the expected commitment to its conclusion. Then, what type 2 solutions would show is that either 1 is not the case, and hence $\neg OB\Gamma$ (because some of the premises are not true) or 2 is not the case, and hence, $\neg O\neg B\delta$ (because the conclusion is, after all, true) and, thus, we do not have to conclude 3 and, consequently, we can accept the validity of the argument. Analogously to the case of type 1 solutions, a proponent of a type 2 solution should be able to explain, ideally, why 1 (2) seems true, which will typically imply abandoning some of the core intuitions governing the concepts involved in the paradox.

This discussion has a somewhat unexpected consequence. In order to adopt a type 2 solution you need to have been confronted with a paradox that fits Definition 1 (the traditional characterization) for, if not, you will be unable to identify the truth of the premises and the falsity of the conclusion as the culprits of your impression that the paradoxical argument does not present the expected commitment to its conclusion. In other words, if you are in front of a paradox for the first time, and the paradox is one of the examples we have seen like Curry's paradox with a true sentence in Curry's sentence, your first impression will be to blame the logic, not the truth value of the truth-bearers involved in the argument. I think this is perfectly reasonable; it would, after all, be very difficult to blame the inductive hypothesis in a Sorites argument that proceeds, say, from true premises to a true conclusion, although we would still have the impression that there is something wrong with the argument.[22]

The second question I wish to address is an objection to Definition 4 that is related to the Preface paradox. The Preface paradox, first introduced by Makinson (1965), asks us to consider an author of an academic book who, in the preface to her book, throws in a caveat to the reader about the errors that the book surely contains. At the same time, though, she is committed to each of the assertions in the book. Thus, on the one hand, she believes that each assertion made in the book, say $a_1, a_2, \ldots, a_n$, is true but, at the same time, given the knowledge of her own fallibility, she also believes that the conjunction of all the assertions in the book is false; that is, $a_1 \wedge a_2 \wedge \ldots \wedge a_n$ is false and, hence, $\neg(a_1 \wedge a_2 \wedge \ldots \wedge a_n)$ is true. This can be represented in the following way (using 'B' for 'the author believes that'):

(i) $Ba_1, Ba_2, \ldots, Ba_n$ (that is because the author believes all her claims in the book to be true)

[21] The principles needed in this inference are (where $\rightarrow$ is the material conditional):

$$O(\alpha \rightarrow \beta) \rightarrow (O\alpha \rightarrow O\beta)$$

$$O\neg\alpha \rightarrow \neg O\alpha,$$

which seem perfectly reasonable, and the fact that $O(B\Gamma \rightsquigarrow B\delta) \rightarrow O(B\Gamma \rightarrow B\delta)$.

[22] See fn. 11.

(ii) $B\neg(a_1 \wedge a_2 \wedge \ldots \wedge a_n)$ (that is because the author is aware of her own fallibility)

And if we accept the principle of agglomeration,

(Agg) $(Ba_1 \wedge Ba_2) \rightarrow B(a_1 \wedge a_2)$

then from (i) we can conclude $B(a_1 \wedge a_2 \wedge \ldots \wedge a_n)$. Hence, the author has inconsistent beliefs; in particular, if we suppose that $B\neg\phi$ implies that $\neg B\phi$ we have a plain inconsistency: $B(a_1 \wedge a_2 \wedge \ldots \wedge a_n) \wedge \neg B(a_1 \wedge a_2 \wedge \ldots \wedge a_n)$.

Consider now, keeping in mind the situation described above, the following argument:

Adjunction Argument

1. $a_1, \ldots a_n$
2. $a_1 \wedge a_2 \wedge \ldots \wedge a_n$,

which seems to be a perfectly harmless argument. According to Definition 4, though, the instance of the Adjunction argument given by the situation described in the Preface paradox will be a paradox; because, in this situation, believing the premises does not commit us, under conditions of epistemic neutrality, to believe the conclusion—that is precisely what the Preface paradox shows; you can believe all the premises while you believe the negation of the conclusion. But even if $a_1, a_2, \ldots, a_n$ are the assertions in the author's book, the resulting instance of the Adjunction Argument is not a paradox, the objector to Definition 4 would claim; it is just a harmless application of adjunction (the principle according to which $\phi, \psi \vdash \phi \wedge \psi$). What this would mean, then, is that Definition 4 is too broad; some arguments that are not paradoxes would be declared as paradoxes.[23]

Notice, though, that the fact that the logical form of an argument seems innocuous does not mean that the argument is. Consider a soritical paradox like Argument 2 in Section 2; its paradoxical status did not depend on its logical form—which was shared by the trivial arithmetical Argument 1 in the same section—but on certain properties of the notions involved in the argument. The case is similar with respect to the instance of the Adjunction Argument where $a_1, a_2, \ldots, a_n$ are the assertions in the author's book. In this case, the argument *is a paradox*, even if its logical form can be instantiated by perfectly sound arguments. Its paradoxical status, though, stems from certain properties of the sentences in the argument, not from its logical form.

The third question I wish to address is the following objection to Definition 4. Given the challenges that have been presented against the normative status of logic,[24] it might seem problematic to characterize the notion of paradox in terms of this very

[23] I am greatly indebted to Elia Zardini on this point.

[24] Above all as a result of the work of Hilbert Harman (see, for example, Harman 1984, 1986, 2002). According to Harman, there is no connection between logic, on the one hand, and norms of reasoning on the other, so that no normative commitment stems from logical validity. For more on Harman's challenge, proposed lines of response to it and more discussion of it see, for instance, Broome (1999), MacFarlane (2004), Field (2009), Field (2015), Russell (2017), and Steinberger (2020) (see also fn. 17). In fact, one line of response to Harman's challenge involves devising principles that articulate a connection between logic and norms of reasoning. Such principles are the bridge principles used at the first part of this section.

same status.[25] To my mind, two responses can be given to this objection. First, note that Definition 4 does not need logic to be normative in order for it to work, but it is enough that logic is *apparently* normative, which is a far weaker claim and a much more plausible one.[26] If logic is apparently normative, we still remain stuck when we are in front of a paradox, for we still have the two conflicting intuitions regarding the commitments that stem from the paradoxical argument: apparently there is no commitment (for the reasons discussed in Section 4) and, apparently, there is (for the apparent normativity of logic). This leaves another possibility open, one that stems from Definition 4 (which did not stem from Definition 1, and hence can be seen as an advantage of the former), and which leads to the second response to the objection: we might interpret pathodoxes like the Curry argument where Curry's sentence is build using 'snow is white' as perfectly sound arguments that do not generate the kind of commitment expected from the point of view of the normativity of logic, so that pathodoxes like these ones would constitute counterexamples to the normativity of logic. In short, Definition 4 can be understood as characterizing the notion of paradox, not in terms of the normative character of logic, but in terms of the *apparent* normative character of logic (which is far less problematic) and, this, in turn, opens a new possible response to the paradoxes: denying that the fact that the paradoxical argument is valid implies that believing its premisses and accepting its validity commits one to its conclusion; to wit, denying the normative status of logic.

One last remark: I think Definition 4 captures well other paradoxes in the field of Philosophy of logic like, for example, the Liar paradox, the paradoxes of material implication or the paradoxes of strict implication, but it remains to see whether this definition can capture other so-called 'paradoxes' in other fields. I will leave this endeavour for future research.

Acknowledgments I am grateful to José Martínez and Elia Zardini for all their extremely helpful comments to previous versions of this paper. I am also grateful to Jose Díez, Oscar Cabaco, Manuel García-Carpintero, Josep Macià, Genoveva Martí, Manuel Pérez Otero, Sven Rosenkranz, Gonçalo Santos, Barbara M. Sattler and a number of anonymous referees. During the writing of this paper, I have benefitted from funding awarded to the project Worlds and Truth Values: Challenges to Formal Semantics (MUNVAL) (2019PIDPID-107667GB-I00) from the Spanish Ministry of Science and Innovation.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declarations

Conflict of Interest The author declares no competing interests.

[25] Thanks to an anonymous referee for very helpful discussion on this issue.

[26] As a matter of fact, the whole discussion around the normative status of logic and Harman's challenge seems to implicitly presuppose that that logic is normative is the default position, which suggests that, at the pretheoretic level, is the most plausible one and, hence, that it is apparently true.

References

Armour-Garb, B. (2017). Introduction. In B. Armour-Garb (Ed.) *Reflections on the liar* (pp. 1–20). Oxford: Oxford University Press.
Barnes, J. (1982). Medicine, Experience and Logic. In J. Barnes, et al. (Eds.) *Science and Speculation*. Cambridge: Cambridge University Press.
Beall, J. C., & van Fraassen, B. (2003). *Possibilities and Paradox*. Oxford: Oxford University Press.
Broome, J. (1999). Normative Requirements. In *Ratio 12* (pp. 398–419).
Broome, J. (2013). *Rationality Through Reasoning*. Oxford: Wiley-Blackwell.
Cave, P. (2009). *This Sentence is False. An Introduction to Philosophical Paradoxes*. New York: Continuum.
Cook, R. T. (2013). *Paradoxes*. Cambridge: Polity Press.
Field, H. (2009). What is the normative role of logic? In *Proceedings of the Aristotelian Society 83* (pp. 251–268).
Field, H. (2015). What is Logical Validity?. In C. Caret, & O. Hjortland (Eds.) *Foundations of Logical Consequence*. Oxford: Oxford University Press. Chapter 2.
Harman, G. (1984). Logic and reasoning. In *Synthese 60.1* (pp. 107–127).
Harman, G. (1986). *Change in view*, (p. 160). Cambridge, Mass: MIT Press.
Harman, G. (2002). Internal Critique: A Logic is Not a Theory of Reasoning and Theory of Reasoning is Not a Logic. In *Handbook of the Logic of Argument and Inference: The Turn Towards the Practical. Volume 1. Studies in Logic and Practical Reasoning* (pp. 171–186). Amsterdam: Elsevier.
Ikuenobe, P. (2014). On the theoretical unification and nature of fallacies. In *Argumentation 18* (pp. 189–211).
Jackson, F. (1984). Petitio and the Purpose of Arguing. In *Pacific Philosophical Quarterly 65* (pp. 26–36). Reprinted in Jac87.
López de Sa, D., & Zardini, E. (2007). Truthmakers, knowledge and paradox. In *Analysis 67.3* (pp. 242–50).
Lycan, W. G. (2010). What, exactly, is a paradox? In *Analysis 70.4* (pp. 615–622).
MacFarlane, J. (2004). In What Sense (if Any) is Logic Normative ofr Thought? ms.
Makinson, D. C. (1965). The paradox of the preface. In *Analysis 25.6* (pp. 205–7).
Oms, S. (2020). La noció de paradoxa. In *Anuari de la Societat Catalana de Filosofia 30–31* (pp. 159–166).
Oms, S., & Zardini, E. (2019a). An introduction to the sorites paradox. In S. Oms, & E. Zardini (Eds.) *The Sorites Paradox* (pp. 3–18). Cambridge: Cambridge University Press.
Oms, S., & Zardini, E. (Eds.). (2019b). *The Sorites Paradox*. Cambridge: Cambridge University Press.
Oms, S., & Zardini, E. (2021). Inclosure and intolerance. In *Notre Dame Journal of Formal Logic, 62 (pp. 201–220)*.
Powers, L. H. (1995). The One Fallacy Theory. In *Informal Logic 17* (pp. 303–314).
Priest, G. (2003). *Beyond the Limits of Thought*, 2nd edn. Cambridge: Oxford University Press.
Priest, G. (2006). *In Contradiction:AStudy of the Transconsistent*. Oxford: Oxford University Press.
Quine, W. V. (1966). The ways of paradox. In W. V. Quine (Ed.) *The Ways of Paradox and Other Essays*. New York: Random House.
Russell, G. (2017). Logic isn't normative. In *Inquiry 63* (pp. 371-388).
Sainsbury, R. M. (2009). *Paradoxes. 2nd*. Cambridge: Cambridge University Press.
Schiffer, S. (2003). *The Things We Mean*. Oxford: Oxford University Press.
Sinnott-Armstrong, W. (2012). Begging the question. In *Australasian journal of philosophy 77,2* (pp. 174–91).
Slaney, J. K. (2010). A Logic for Vagueness. In *Australasian Journal of Logic 8* (pp. 100–134).
Sorensen, R. (2003). A Brief History of the Paradox. Oxford: Oxford University Press.

Steinberger, F. (2020). The normative status of logic. In *The stanford encyclopedia of philosophy, Edited by Edward N. Zalta. Winter 2020. Metaphysics Research Lab, Stanford University*.
Walton, D. N. (1991). *Begging the Question: Circular Reasoning as a Tactic of Argumentation*. New York: Greenwood Press.
Walton, D. N. (1989). *Informal Logic*. Cambridge: Cambridge University Press.
Walton, D. N. (1995). *A Pragmatic Theory of Fallacy*. Tuscaloosa: University of Alabama Press.
Wright, C. (1975). On the coherence of vague predicates. In *Synthese 30* (pp. 325–65).
Zardini, E. (2008). A model of tolerance. In *Studia Logica 90* (pp. 337–368).
Zardini, E. (2021). Substructural approaches to paradox: an introduction to the special issue. In *Synthese 199* (pp. 493–525).