AUSTRALASIAN JOURNAL OF LOGIC

The Paradox Paradox Non-Paradox and Conjunction Fallacy Non-Fallacy

Noah Greenstein*
me@noahgreenstein.com

Abstract

Brock and Glasgow recently introduced a new definition of paradox and argue that this conception of paradox itself leads to paradox, the so-called Paradox Paradox. I show that they beg the question during the course of their argument, but, more importantly, do so in a philosophically interesting way: it reveals a counterexample to the equivalence between being a logical truth and having a probability of one. This has consequences regarding norms of rationality, undermining the grounds for the Conjunction Fallacy.

1 Introduction

Brock and Glasgow [2022] recently introduced an excellent definition of paradox. Perhaps, to make a splash, they then use this definition to generate a new paradox, their so-called Paradox Paradox. Unfortunately, for them, they beg the question against themselves during the course of their argument. Fortunately, for us, this reveals further insight into the relation between logic and probability: paradox stands as a counterexample to the fundamental presupposition that truth is equivalent to a probability of one. Specifically, paradoxes are things that have apparently true constituent propositions while still being false, whereas the total probability must always equal the sum of its parts. So even though we may combine all the different options to give a probability of one, if this is a paradoxical statement, it will be false. Once truth and probability can so come apart, probabilistic norms

*For my father.

of rationality become undermined, including the justification for the Conjunction Fallacy.

Note, though, that while what follows will be structured as a kind of "response" paper, the overall point will stand more generally: the Paradox Paradox is merely a convenient counterexample, but one that shows a way to make logic, probability, and rationality diverge.

2 The 'Paradox Paradox' Non-Paradox

Brock and Glasgow [2022, p. 1] introduce these three seemingly acceptable premises:

The Ontological Thesis: There are paradoxes.

The Conceptual Thesis: A paradox is a set of inconsistent propositions such that it is rationally permissible to find the conjunction of any proper subset of them plausible even while being aware that the set as a whole is inconsistent.

The Normative Thesis: If A entails B, then it is rationally impermissible for an agent aware of the entailment to find A more plausible than B.

From these premises they derive a paradoxical conclusion, that there both are and are not paradoxes. However, I will now show that their argument fails.

First, some preparation: Both they and I assume that the Ontological Thesis is uncontroversial. However, their notion of 'plausibility' will need to be clarified in the other theses. Let's start with the Normative Thesis.

From a logical point of view, if A entails B, then the conclusion B follows from the premise A (assumed to be true) and the conclusion should have the same truth-value as the premises. Formal logic is generally taken to be truth-preserving. So it isn't that the premise cannot be more plausible than the conclusion: the premises have to be *exactly as plausible* as the conclusions, as they are all equally *true*. Hence the way the Normative Thesis is phrased, so that the premise could be less plausible than the conclusion, is technically imprecise.

I mention this imprecision for two reasons. First, it may seem inconsequential now and for most of the argument. It will not yield any obvious contradictions. However, it is a symptom of a much deeper problem, and will be referred to throughout the paper.

Second, it is a convenient entry point into discussing the more specific issues. I will proceed in this section by first clarifying the theses of Brock and Glasgow with respect to it and then re-evaluating their argument.

The imprecision raises a doubt whether the 'plausibility' above has to do with formal logic. Since they "take this constraint on rational belief [The Normative Thesis] to be nearly self-evident" [Brock and Glasgow, 2022, p. 4] it seems like they could be discussing logic. They then also appeal to the "Conjunction Fallacy" of Tversky and Kahneman [Brock and Glasgow, 2022, p. 4], but this appeal is merely to point out that Tversky and Kahneman also argue that it is irrational to find the premises more plausible than the conclusions.

However, the Conjunction Fallacy refers to a *probabilistic* norm [Tversky and Kahneman, 1982, p. 91], not one of formal logic. The Conjunction Fallacy (or Rule) holds that the likelihood of any one conjunct of a conjunction is at least as likely as the entire conjunction, $P(A\&B) \leq P(B)$. That people often do not follow this norm of probability accounts for the name: Conjunction *Fallacy*.

Perhaps this is closer to what Brock and Glasgow have in mind with the Normative Thesis. If they are thinking probabilistically, then it makes sense that the phrasing of the Normative Thesis follows the structure of the probabilistic norm $P(A\&B) \leq P(B)$: saying the premise cannot be greater than in plausibility than the conclusion is equivalent to saying the premise is less than or equal to the conclusion's plausibility. They also appeal to subjective probabilities when deriving their contradiction, "... if the agent finds r plausible, then the subjective probability of r for the agent is greater than 0.5" [Brock and Glasgow, 2022, p. 5]. Hence it appears that they are cashing out the idea of plausibility as having a subjective probability greater than 0.5.

This gives us:

The Normative Thesis*: If A entails B, then it is rationally impermissible for an agent aware of the entailment to find A more *subjectively probable* than B.

With this clarification of Normative Thesis, we can also update the Conceptual Thesis:

The Conceptual Thesis*: A paradox is a set of inconsistent propositions such that it is rationally permissible to find the conjunction of any proper subset of them *more subjectively likely than not* even while being aware that the set as a whole is inconsistent.

Conceptual Thesis* captures the idea that all the statements in a paradox can seem subjectively likely by themselves, but when put together result in inconsistency.

Now let us turn to their derivation of an inconsistency, that there are and are not paradoxes [Brock and Glasgow, 2022, p. 5]:

> Consider any arbitrary set of inconsistent propositions S. Next arbitrarily separate any one of its constituent propositions p from the conjunction of the remaining propositions r. Either (a) rational agents can ascertain that S is inconsistent and find r plausible, or (b) they can't. Suppose (a) they can't ascertain that S is inconsistent and find r plausible. In that case, S is not a paradox (from the Conceptual Thesis).
>
> Suppose, instead, that (b) a rational agent can ascertain that S is inconsistent and also find r plausible. In that case, if the agent finds r plausible, then the subjective probability of r for the agent is greater than 0.5. After all, if the subjective probability of r were lower than 0.5, a rational agent would find its negation more plausible. And if S is inconsistent, and if the agent is aware of this, she will see that r entails not-p. And so the subjective probability for the agent of not-p will also be greater than 0.5 (from the Normative Thesis). Therefore, the subjective probability for the agent of p will be less than or equal to 0.5, which means the agent will not find p plausible. But, then, since S includes p, S is not a paradox (from the Conceptual Thesis).
>
> Either way, then, S is not a paradox. And because S is any arbitrary set of inconsistent propositions, there are no paradoxes. But there are paradoxes (from the Ontological Thesis). That's paradoxical!

The issue here is that their just-defined Conceptual Thesis explicitly states that paradoxes are inconsistencies where all the proper subsets of propositions are plausible, that is, subjectively more likely than not. Let us take them at their word and assume this to be our working definition of paradox. Given the Ontological Thesis that guarantees that paradoxes exist, what should we infer from some arbitrary inconsistency S? *A priori*, there are two cases: either S is a paradox or else it is not:

1. If S is a paradox, then, by assumption, we could find some constituent proposition r plausible. But, since we assume that S is a paradox, it follows from the definition that the rest of the constituent propositions p are also plausible. Hence it would *not* follow that not-p will have a subjective probability greater than 0.5.

2. On the other hand, consider if S is not a paradox but just an inconsistency, and we also find r plausible. Then it will follow that the rest of the constituent propositions p should not be plausible. Hence, in this case, it *would* follow that not-p will have a subjective probability greater than 0.5.

Only in Case 2, where we assume S is not paradoxical, can it be argued, via looking at the subjective probabilities of the constituent propositions, that S is not a paradox. This is, of course, *not interesting*: it is the expected result when we assume S is not a paradox that our analysis returns that it is not a paradox.

Case 1, also returned the expected result, that a paradox is analyzed as an inconsistency with plausible constituents, which accords with their definition of paradox: the Conceptual Thesis.

However, they claimed in the above quote, that even when S is a paradox that "a rational agent would find [some constituent proposition's] negation more plausible". Since their own account of paradox states that any subset of constituent propositions of a paradox will be plausible, it follows that their negations are *not* more plausible. That they then say a rational agent would not even countenance this possibility, immediately assuming any inconsistency to have some less than plausible constituent propositions, is very strange: it begs the question against Case 1 — against their own account! — where S is paradoxical according to the Conceptual Thesis.

Regardless of the details of these underlying issues, getting the uninteresting, expected result when analyzing paradoxes is a major problem for Brock and Glasgow in terms of the Paradox Paradox. For the Paradox Paradox to be a paradox, it needs to have the unexpected, paradoxical result that there are no paradoxes. They claimed that any arbitrary inconsistent S could not be paradoxical, but, following their own account of paradox did yield the correct result. Therefore, their argument that there are no paradoxes fails.

To lend support to this conclusion, recall the nature of my analysis above: this started with a clarification of their concepts, giving them more precise definitions, and then a re-evaluation of their argument. No counter-examples or alternative theories were presented. All three of their premises were *granted*, at least in modified form.

Then I tried to derive the Paradox Paradox using these more precise versions of their theses: the paradox's existence was not assumed one way or the other. However, with the greater precision, the needed paradoxical conclusion could no longer be derived via Brock and Glasgow's argument. Hence, their argument should not have worked in the first place and no subsequent iterations — without major changes to their initial theses or their argumentation — will help.[1]

3 The 'Conjunction Fallacy' Non-Fallacy

Although the previous section demonstrated *that* the argument for the Paradox Paradox fails, it doesn't explain exactly *why* it failed. It was suggested, though, that there was something imprecise about the above account that had to do with a shift between logical reasoning and probabilistic reasoning. This is significant because if Brock and Glasgow were able to invoke the probabilistic reasoning, their argument would go through: the result of the previous section (Case 1) prevented that step from occurring. Hence, whatever the difference is between logical reasoning and probabilistic reasoning, it is *substantive*, contrary to the general belief that these kinds of reasonings can be used interchangeably.

For instance, we can represent the relation between probability and truth as:

$$(P(x) = 1) \leftrightarrow (Tr\langle x\rangle = \top)$$

This says that if something has probability equaled to 1, if and only if it is alethically evaluated to True. Often enough the symbols $\top$ and $\perp$ aren't used for True and False but are replaced by 1 and 0, yielding formulas like: $Tr\langle x\rangle = 1$. That it is common practice to switch back and forth between these formalisms reiterates how accepted the equivalence between the two concepts are, and supports Brock and Glasgow [2022, p. 4] in claiming self-evidence for their use of probabilistic reasoning.

As argued above, however, their theory of paradox provides a counterexample to this interchangeability. So what sets paradox apart?

To answer this, it is instructive to look at the argument for the Paradox Paradox inconsistency. Brock and Glasgow [2022, p. 5] begin "Consider any arbitrary set of inconsistent propositions S. Next arbitrarily separate any one of its constituent propositions p from the conjunction of the remaining propositions r." Note the *arbitrary separation* of one of the propositions from the rest. It is this move that runs contrary to the spirit of the Conceptual Thesis in that paradoxes only exist for the *set of propositions as a whole*. That is, every proper subset is related to the others by being more subjectively likely than not *while the set as a whole is inconsistent*. Hence there is a holism at work in paradox, and separating out one of the constituent propositions breaks this holism.

[1]Thanks to a reviewer for helping clarify this point.

The problem is that the holism is not represented by the logic given in the argument. A set or conjunction of propositions is not a holistic representation because it can be trivially decomposed into its constituent propositions. This disregards the inter-related structure of the Conceptual Thesis in that when we look at the whole set, then each of the propositions, or any proper subset combination of them, will be subjectively more likely than not. So, if we want to describe paradox using the Conceptual Thesis, there has to be more to the logic than a mere *conjunction* of propositions.

Realizing this sophisticated structure in logic might be possible. However, as Brock and Glasgow point out, their account of paradox is not universally applicable: it only works for countable sets [Brock and Glasgow, 2022, footnote 1] and there are phenomena that we think of as paradoxes, like the Sorites, that don't *prima facie* fit the form of the Conceptual Thesis [Brock and Glasgow, 2022, p. 6]. Moreover, investigating different ways to represent the paradoxical holism — or potential *holisms* — in formal logic, surveying different logics suited to this task, and determining if there is a single best logical schematization, would go well beyond my current scope.[2]

The goal was, recall, to show the difference between the logical reasoning and the probabilistic reasoning, and we now have an answer. In the probabilistic reasoning used, it is assumed that the probabilities of the constituent propositions are independent of each other. However, this clashes with the holism of paradoxes: if we want to capture the paradoxical phenomenon, then the paradox's constituent propositions can only be analyzed as an inter-related group. Hence these probabilistic norms will not apply in such situations.

The wider issue is that it is not clear when and where holisms and holistic reasoning arise. Paradoxes, almost by definition, show up in times and places when we are not expecting them. Also, as just mentioned above, there are edge cases to the Conceptual Thesis, that is, paradox-like phenomena that may not strictly fit the Conceptual Thesis, but act similarly. Hence, there may be even more instances of these sorts of holisms that have gone unnoticed or, at least, underappreciated, some of which will be discussed in the conclusion below.

This brings us back to the Conjunction Rule, $P(A\&B) \leq P(B)$ [Tversky and Kahneman, 1982, p. 90], which, recall was given as support for the Normative

[2] While such a formal investigation would take us too far afield, a first thought is to look at something like Independence Friendly Logic, which allows for non-compositionality [Hintikka and Sandu 2011, p. 423; Tulenheimo 2022, §4.4], and hence could realize a structure that would not be trivially reducible. Though appealing to non-compositionality may be overkill, this shows that there is potentially at least one way to formally represent a holism.

Thesis. The Conjunction Rule does not countenance holistic exceptions. This means that there may be instances when people are not following the Conjunction Rule, that is, apparently committing the Conjunction Fallacy, when they are reasoning rationally, albeit in a holistic way. Hence the Conjunction Rule/Fallacy will be inapplicable in those instances.

To be clear: I am not saying that all instances of the Conjunction Fallacy are wrong, or even that most are. What I am saying is that it is based on a limited conception of logic — as this analysis of paradox has shown — which means that the Conjunction Rule cannot be generally applied as a norm of rationality. People could be violating the norm when it does apply, that is, when they ought be treating the conjuncts as separate events. However, since we haven't been looking for holistic logical phenomena or reasoning surrounding holistic phenomena, it is therefore premature to declare agents who commit the Conjunction Fallacy to be irrational.

Consider the famous "Linda Problem" example:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.

Tversky and Kahneman [1982, p. 92] conducted experiments asking participants to rank the probability of various statements about Linda after reading the above information. These were two of those statements:

(BT) Linda is a bank teller

(BT & F) Linda is a bank teller and is active in the feminist movement.

The vast majority, between 85% and 89% of subjects, placed the conjunction BT&F above the singleton BT as the more probable statement [Tversky and Kahneman, 1982, p. 93] and hence are not following the Conjunction Rule.

They interpreted this result as a "Framing" issue, that is, how the question was presented or "framed" caused people to diverge from the Conjunction Rule. In contrast, I am suggesting that people are *not acting irrationally* in these cases because people are interpreting the problem holistically, as a case where the Conjunction Rule does not apply: people are reasoning about Linda as a whole person and not just as group of properties that can each be analyzed independently of the others.[3] If we do not reduce Linda to some group of independent, unrelated properties — which is a *morally* commendable position — then it makes less sense to speak about those individual properties and their probabilities. Hence holistic thinkers rank the single conjunct BT as less likely than the whole conjunction, and there is no fallacy of rationality here.

4 Conclusion

I argued that the Conceptual Thesis of Brock and Glasgow postulates a philosophically important holistic structure holding between the propositions that constitute a paradox. This structure cannot not be reduced to a simple logical conjunction, which sets it apart from other inconsistencies, and, at minimum, also means that certain common appeals to probabilistic reasoning are no longer valid. Therefore paradox stands as a counterexample to the equivalence we have assumed held between logical truth and probability.

However, even if one does not agree with the above analysis of paradox, the disconnect between logical truth and probability remains. Say, for the sake of argument, we take quantum entanglement as a kind of holism where the entangled particles fundamentally cannot be represented independently from each other. Then quantum entanglement would stand as a counterexample where mere conjunction provides an insufficient description of the world. Our logic should be able to reflect this.

So, given some such holistic phenomenon, it would provide a counterexample to the equivalence of logical truth and probability — in the place of, or in addition to, paradox. But, *it does not even matter that such an example exists in physical reality*: as the quantum entanglement example presents a cogent picture of the world, even if physics does not turn out to be this way, our logic needs to be able account for such a situation. Hence, that the holistic account *is at least possible*, if not plausible, demonstrates that logical truth and probability can come apart in the way described above.

On this line of thought, let's consider another example: consciousness. Chalmers defined two classes of problems associated with consciousness, the 'Easy

[3] I agree that the way the question is presented may effect the likelihood of people reasoning holistically. So nothing argued here should be taken as directed against the framing effect. Also, Tversky and Kahneman [1982, p. 93] found people ranked $P(F) > P(BT\&F) > P(BT)$. Since only the feminism was mentioned in the original story, it makes sense that this property is a locus of the holism, that from which everything else cannot be independently discussed.

Problems' and the 'Hard Problem' [Chalmers, 1995, ppg. 200-201]. The Easy Problems of consciousness include explaining

1. the ability to discriminate, categorize, and react to environmental stimuli;

2. the integration of information by a cognitive system;

3. the reportability of mental states;

   et cetera.

The Hard Problem is to explain subjective experience, what it is like to be conscious. The claim is that even if we can scientifically explain all the individual 'Easy' problems of consciousness, we would still not have an explanation of our experience of consciousness.

Putting this in my terms is a stretch, but not a big one: what it is like to be conscious (The Hard Problem) cannot be reduced to its constituent pieces (The Easy Problems). So described, consciousness represents a kind of holism. Now, my point is *not* that consciousness is accurately described by the Hard and Easy Problems. Instead, the point is that *philosophers describe consciousness in these terms*: the Hard Problem of consciousness is part of the standard literature in philosophy of mind. Hence there is a large, ongoing discussion that characterizes consciousness in the holistic way given by this paper.

So, insofar as our logic needs to be able to accommodate the *discussion* of the Hard Problem of consciousness, our logic must be able to contend with holisms. Therefore, this again demonstrates that logical truth and probability can come apart in the way described above.

Now, someone may respond by saying that these are problems to be solved, and then, once solved, they won't be problems any more. Hence there won't be any further talk of holisms.

I grant these problems may yet be solved. However, even if they are, it matters not. First, insisting that all holisms are problems that will eventually be solved is dogmatic, and begs the question against this account. Second, that we can discuss a holistic analysis in an unproblematic way, as we have been doing here and as I claimed happens in the philosophy of mind, shows that there is linguistic practice surrounding holisms. Hence holisms already exist within our linguistic repertoire, and it would be perverse to ban them from our discourse or logic.

Another objection is that my account is far too strong: if paradoxes only exist for the set of propositions as a whole, this runs contrary to generally accepted understanding that the probability of some event has something to do with its priors, that is, its constituents.[4] In one respect, I am willing to bite this bullet. Paradoxes are strange things, and having them be unique, as compared to other phenomena with respect to probability, does not seem all that problematic. It would just be one more unique property of an already unique phenomenon. Hence, insofar as the holism claim only applies to paradoxes and the few other similar phenomena, then, so be it, they do not follow the general pattern with respect to the probability of their priors.

More can be said, though. First, recall the previous section where it was noted that paradoxes quite often take us by surprise. We thought we were on solid logical ground, and then they pull the rug out from under us. This is empirical evidence that the priors in paradoxes do not act like priors for other events. That is, paradoxes surprise us because their priors do not act in the expected way, and it is our experience with paradoxes that demonstrates this. Although this is anecdotal, it supports the claim that paradoxes are different from other phenomena when it comes to probabilistic priors. Thus, regardless of my overall argument here, there is still good reason to think paradoxes do not follow the general pattern with respect to probabilistic priors.

Also recall that the Conceptual Thesis of Brock and Glasgow [2022, p. 1] states: A paradox is a set of inconsistent propositions such that it is rationally permissible to find the conjunction of any proper subset of them plausible even while being aware that the set as a whole is inconsistent. Although my final conclusion here is wider than Brock and Glasgow's account of paradox, they did specifically formalize paradoxes as things that do not follow the probability of their constituents propositions, that is, their priors. Hence, I believe we are in agreement on this point, so there are a least a few of us that are willing to state, in print, that paradoxes do not follow the typical pattern with respect to probability.

Lastly, consider the priors of a fundamentally probabilistic quantum phenomenon, such as the radioactive decay of uranium into lead. We know the probabilistic rate at which the decay occurs. But, for any individual block of metal, that is all we know: this block may decay twice as slow, while another decays twice as fast. There are no more priors to be known about any individual block, only that, on average, they decay at a certain rate. My point is that we *already allow for exceptions* to our understanding of priors in probability. Radioactive decay, wave-function collapse, the uncertainty principle, or any other of a litany of quantum phenomena have restricted or no priors. And, if we are already willing to accept quantum exceptions to the usual pattern of probabilistic reasoning, then I am

[4]Thanks to a reviewer for raising this issue.

merely suggesting we add a few more: the limited number of holistic phenomena. Hence it is not so much that we are biting a bullet in changing our understanding of priors in probability — we have already done that — but expanding the range of phenomena that do not follow the general pattern.

Initiated by an imprecision in the account of logical entailment, these holistic examples and theories, whether holism ultimately applies to paradox or not, witness the failure of the equivalence of logical truth and probability. Beyond deepening our understanding of the relationship between logic and probability, the consequences of this analysis reach beyond philosophy: widely accepted norms of rationality, such as psychology's Conjunction Fallacy, are undermined. Therefore, the holistic analysis of the Conceptual Thesis represents an advance not only in our understanding of paradoxes, but of logic and rationality as well.

References

Stuart Brock and Joshua Glasgow. The paradox paradox. *Synthese*, 200(2):1–7, 2022.

David Chalmers. Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3):200–19, 1995.

Jaakko Hintikka and Gabriel Sandu. Game-theoretic semantics. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*, pages 415–465. MIT Press, 2 edition, 2011.

Tero Tulenheimo. Independence Friendly Logic. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2022 edition, 2022.

Amos Tversky and Daniel Kahneman. Judgments of and by representativeness. In Daniel Kahneman, Paul Slovic, and Amos Tversky, editors, *Judgment Under Uncertainty: Heuristics and Biases*, pages 84–98. Cambridge University Press, 1982.