

1 Introduction

Most philosophers agree that outside of the formal parts of philosophy, we cannot offer rigorous *knockdown proofs* of philosophically substantive theses. We might be able to prove things in formal logic, mathematics, and decision theory, and those results can be philosophically important. But knockdown proofs of philosophically substantive theses in regular, nonformal philosophy? Ha ha . . . nope. Most of us without hubris abandon the goal of giving knock-down (informal) proofs of philosophically substantive conclusions.

As a result, many of us retreat to offering reasonable, “interesting,” yet hardly foolproof arguments for theses of the form “This solution to the problem is true/false.” If one has struggled for a number of years with such projects, then one will, unless wildly confident in one’s abilities, have moments of despair about even this more modest goal. After all, you will, if you are sufficiently reflective and honest with yourself, eventually see that there always are good (yet inconclusive) objections to your arguments that you cannot block. So, we often *retreat once again*, watering down our conclusions to “This view has such-and-such going for it” or “This criticism of such-and-such view is no good – unless, of course, the criticism gets changed radically in order to maneuver around the problems I’ve detailed here.”

Another disappointment is that most philosophers now admit that philosophy, by itself, is not nearly as strong as science when it comes to showing how common sense is mistaken. In the olden days, philosophers advanced bold arguments purporting to prove that many of the average idiot’s beliefs were laughably false, to put the point boldly and unkindly. Now we hang our heads and mumble that although science is up to the task, our profession is too weak to deliver the goods. Some of us even feel intellectually tough in admitting this loudly, without the mumbling. It’s akin to standing up and saying, “Yes, I am a sinner! I am! I admit it!”

So, conventional wisdom says “No knockdown arguments in normal philosophy” and “No refutations of common sense from philosophy.” What would be really crazy would be philosophical arguments that are *both* knockdown *and* refutations of common sense. That would refute both parts of conventional wisdom with one blow.

That’s the way things seemed to me – until recently. Now I suspect, but certainly do not believe, that we can give knockdown proofs of highly counterintuitive theses, refuting both parts of conventional wisdom with one blow. However, we have to do it in a new way, starting from *exhaustive analyses of philosophical problems*. Bizarrely enough, we may be able to prove philosophically substantive theses without taking any stand on the solutions to any of those problems we are analyzing.

To get an idea how this might work, pretend that each of the following three claims regarding some intellectual problem is highly intuitive: C_1 , C_2 , $\sim(C_1 \& C_2)$. It's child's play to derive a contradiction from them. Your first reaction to such a situation is to conclude that at least one of the three claims is false. That may seem a safe inference to make. However, if you decide to be more cautious, to look for proofs instead of merely persuasive arguments, you will realize that such a conclusion is hasty. For one thing, maybe some contradictions are true. For another thing, maybe there is some subtle equivocation in the trio of claims that blocks the application of the derivation of the contradiction (so they don't really have the logical forms they appear to have). For yet a third thing, maybe elementary logic is flawed in some deep fashion, so the obvious derivation applies to the three claims but the inferences aren't truth-preserving.

Having such a cautious attitude doesn't preclude one from confidently drawing conclusions, however. We can be *extremely confident* in concluding that the truth lies in one of the following four possibilities:

- one or more of the three claims C_1 , C_2 , and $\sim(C_1 \& C_2)$ isn't true
- the trio of claims is true but there is some subtle equivocation or other linguistic difficulty present in them that makes the derivation of the contradiction not apply, so no contradiction results from the trio
- the trio is true and there is no such linguistic difficulty (so the derivation does apply just as we expected) but one of the elementary sentential inference rules in the derivation isn't truth-preserving, so no contradiction results
- the three claims are true, the derivation goes through as expected, the inference rules are truth-preserving, and thus a contradiction is true too.

If we realize that the claim "An equivocation or similar linguistic complication was present that ruined the obvious application of the derivation" is philosophically counterintuitive in the rough sense that its truth would require a huge change in our beliefs, then we can safely conclude that *whatever the truth is about the intellectual problem in question*, it's philosophically counterintuitive, since, as we just saw, there are exactly four options and each has that quality. Therefore, our analysis of the problem itself, without taking any stand on its solution, has yielded two interesting results: one, the disjunction of the four claims is true; two, a philosophical argument proves that there exists a philosophically counterintuitive truth (viz. one of the disjuncts, although we don't know which one). In this Element, I show that something akin to this scenario applies to many (not all) philosophical problems. As we will see, its consequences are philosophically significant.

I have long had the gut feeling that many philosophical paradoxes are amazing in the sense that they *force* one to adopt a view that is firmly, even wildly, against common sense. I don't think I'll ever have a decent reason to think that such-and-such a detailed response to a given paradox is correct. For instance, by my lights the presently existing public considerations regarding the Sorites paradox are in favor of epistemicism. But why on earth would I think that those considerations – the ones we have *today*, located in the *public* sphere – are good enough to point us in the right direction? I can't think of any good reason to think that; in fact, there are excellent reasons (not given here) to think that that's just plain false. In any case, I wondered whether I could justify my gut feeling that certain philosophical paradoxes refute portions of common sense – no matter what the correct solution was to those paradoxes. So, I tried to find that justification, with reference to several philosophical paradoxes. Like many philosophers familiar with paradoxes, I used to think of paradoxes as seemingly valid arguments from commonsensical premises to highly counterintuitive conclusions. So, I figured that the mere existence of the paradoxes proved that either a commonsensical premise is false, a seemingly valid argument is invalid, or a highly counterintuitive conclusion is true. As hinted at with the $\{C_1, C_2, \sim(C_1 \ \& \ C_2)\}$ scenario discussed earlier in this Introduction, now I see that that diagnosis misses some possibilities.

What the paradoxes have in common is that an exhaustive analysis of each of them reveals that any of its proposed solutions is philosophically counterintuitive, in a sense to be defined in Section 2. That's a nice thesis, but things get much more interesting when we pool these analyses together. When we do so, we find interesting metaphilosophical and epistemological insights on six topics:

- philosophical progress
- agreement in philosophy
- knockdown arguments in philosophy
- the wisdom of philosophical belief
- the epistemic status of metaphysics
- the power of philosophy to refute common sense

In Section 2, I briefly elaborate on those six topics. But if you are impatient to hear the *central* take-home lesson, which lies in the intersection of epistemology and metaphilosophy, here it is:

As I have pointed out, many philosophers think that philosophy never refutes common sense; many philosophers also think there are no knockdown arguments in nonformal philosophy. What would be extraordinary would be philosophical arguments that did both! However, that's what I argue for here: There

are many arguments that (1) clearly belong to nonformal philosophy, (2) are knockdown, and (3) conclude that some violently counterintuitive claims are true; furthermore, once the typical philosopher is aware of the truth of that thesis, she should, in order to be wise, withhold judgment on a colossal number of claims, even highly commonsensical ones. Hence, a lesson of philosophy's success is that once the wise philosopher becomes aware of these facts, she will suspend judgment much more often than she once did. In that sense, we fall doxastic victim to our own argumentative success.

2 Our Six Topics

Philosophers have been complaining about metaphysics for a long time, claiming that it is bullshit when compared to most other areas of philosophy (e.g. Ladyman and Ross 2007). Even today, almost all of us encounter jabs at it, in conversation and on social media if not in print. Up until now, the metaphysician's best response has been: "Oh yeah? Let's see your response to such-and-such metaphysical problem, if you think metaphysics is bullshit." The critics and defenders are rarely impressed with the ensuing discussion, should there be any at all.

Philosophers have been complaining about the lack of philosophical progress for a long time (e.g. some of the papers in Blackford and Broderick 2020). Even today, almost all of us encounter jabs at it, in conversation and on social media if not in print. Most complainers are willing to admit that there are some forms of philosophical progress: new distinctions are discovered that are philosophically key, new problems are discovered, new theories are formulated, new arguments are constructed, new thought experiments are conceived, new fields are generated, and so on. These new things help us see deeper into certain topics, in ways that are hard to articulate non-metaphorically. But even if we grant all of those forms of progress, philosophical progress is anemic compared to scientific progress when it comes to getting substantive, positive truths that are answers to burning questions. It's not like we have solved many really big problems. If we had, then why the hell would we still be reading Aristotle and Kant for solutions instead of as historians only?

Philosophers have been complaining about the power of philosophy to refute common sense for a long time (e.g. Moore 1925; Lewis 1973; Fine 2001; Lycan 2001, 2019; Gupta 2006; Kelly 2008; Schaffer 2009). Even today, almost all of us encounter jabs at it, in conversation and on social media if not in print. These philosophers admit that science is up to the task, and some admit that formal philosophy is as well. The more modest ones are willing to admit that maybe, just maybe, philosophy can be transformed in some surprising way such that

nonformal philosophy refutes common sense in the future. But up until now? No way. They think that even if Moore's response to the argument for skepticism was flawed, the *Moorean move* is reasonable when one is confronted with a philosophical attempt to refute common sense:

One starts out believing P, where one is fully aware that P is utterly commonsensical. One next encounters what one recognizes to be a nonformal philosophical argument A against P. In response, one forms the belief B that *P is significantly more warranted than at least one of the premises of A*. One then retains P with little or no change in confidence even if one admits that one has not (yet) found any flaw in A. This retaining of P is done on the basis of B.

Philosophers have been complaining about the lack of agreement in philosophy regarding philosophically substantive claims for a long time (e.g. van Inwagen 2006; cf. Frances 2017). Even today, almost all of us encounter jabs at it, in conversation and on social media if not in print. All one needs to do is scan the results of the PhilPapers surveys to witness how we fail to reach agreement on just about anything philosophically substantive. This is utterly different from what we find in the sciences.

Philosophers have been complaining about the lack of knockdown (and nonformal) philosophical arguments for philosophically substantive claims for a long time (e.g. Lewis 1983; van Inwagen 2014; cf. Ballantyne 2014; Keller 2015; McGrath and Kelly 2017). Even today, almost all of us encounter jabs at it, in conversation and on social media if not in print. For one thing, if such arguments existed, there certainly wouldn't be so much disagreement on philosophically substantive claims. But, again, just look at the results of the PhilPapers surveys.

As much as philosophers like to complain about the comparative inadequacies of metaphysics, the failure of nonformal philosophy to refute common sense, the anemic nature of philosophical progress compared to that of science, the lack of agreement on substantive philosophical matters, and the failure to produce knockdown nonformal philosophical arguments for philosophically substantive claims, if my experience is at all representative, then what philosophers *really hate* is the idea that the typical wise philosopher should, epistemically if not professionally, suspend judgment on philosophical claims. Think about it: You work for years and years defending your niche position, defending it in multiple publications, and you're not epistemically allowed to even believe it? Are you serious?

What this means is that my theses in this Element will be resisted. I'm reminded of Franklin Roosevelt's address announcing the Second New Deal on October 31, 1936: "Never before in all our history have these forces been so

united against one candidate as they stand today. They are unanimous in their hate for me – and I welcome their hatred.” I’m joking.

In the Introduction, I asserted that we can prove philosophically counterintuitive results from analyses of philosophical problems – and we can do it without taking any stand on the solutions to any of those problems. In order to do this, we start by giving a logically exhaustive analysis of several traditional philosophical paradoxes. The analysis allows us to prove disjunctions that have a small number of disjuncts and have the following features:

- The arguments for the disjunctions are “knockdown” arguments, pretty much however one wants to precisify that notion in a reasonable manner.
- We philosophers are strongly disposed, after seeing the proofs, to agree that the disjunctions are true.
- Each disjunct is *philosophically counterintuitive*: If it is true, then a great many of our ordinary commonsensical beliefs and/or a significant portion of our most confidently held ordinary beliefs or belief-dispositions are false, or key philosophical ideas held by a large portion of philosophers are false. (More on this characterization in Section 4.)
- Once we philosophers are aware of the truth of the disjunctions, we realize that we have been *wildly* wrong about language, logic, truth, or ordinary empirical matters.
- Awareness of the truth of the disjunctions makes the typical wise philosopher suspect that she should not trust her judgment, in a profound manner.
- Once one is aware of the truth of the disjunctions, the typical philosopher must, in order to be wise, suspend judgment on an enormous number of claims, even many of the most certain ones.

As a bonus, the materials used to prove the disjunctions cast light on whether metaphysics is bullshit (it’s not; in fact, in one key respect it is superior to some other areas of philosophy), what kind of substantive philosophical progress there is on particular philosophical claims (it exists, although in an unexpected form), and whether (nonformal) philosophy – instead of science – can refute common sense (yes).

3 How to Analyze a Philosophical Problem: The Sorites

Philosophical problems that stick around for centuries often fall into one of two classes. First, there are those that have multiple proposed solutions that are not *terribly* counterintuitive but we collectively have not figured out which solution is right. Second, there are the really hard problems: those that seem to *require*

a highly counterintuitive solution but we haven't figured out which is true. Opinions will differ about which class a particular problem is a member of; relatedly, there will be loads of borderline cases. Despite those issues, there are obvious citizens of the second class. Aficionados of those problems usually take it for granted that any mature response to them will include a claim that is not mildly odd but highly counterintuitive.

This isn't a book on the Sorites paradox, even though I am going to analyze it. What is relevant about the paradox is the fact that, like many philosophical paradoxes, it can be used to prove what I call a *doxastically distressing disjunction*. To that end, consider the following SC claims ("S" for "sorites," "C" for "claim").

SC₁: Anyone worth less than \$1 (US) is (financially) poor.

SC₂: Either it's not the case that anyone worth less than \$1 is poor, or anyone worth less than \$2 is poor. (In other words, if everyone with less than \$1 is poor, then so is everyone with less than \$2.)¹

SC₃: Either it's not the case that anyone worth less than \$2 is poor, or anyone worth less than \$3 is poor.

...

SC_{LAST}: It's not the case that anyone worth less than \$10¹² is poor.

On the face of it, the SCs are collectively inconsistent, since one can easily derive \sim SC_{LAST} from the other SCs using simple inferences (more on that in this section).

This Element is not concerned with the solutions to Sorites or any other classic philosophical problem. I don't care what your favored solution is, I don't want to hear you blather on about it, and it won't matter to my arguments at all. (The harsh language, which will be repeated in what follows, is for both amusement and, more importantly, hard emphasis that the solutions to the paradoxes do not affect the arguments of this Element.) Instead, we are focusing on the philosophically significant consequences of these problems themselves, not specific proposed solutions to them.

There are exactly five possibilities with regard to the SCs: The first three collectively cover all the ways the conjunction of the SCs can be false and the last two cover the two ways it can be true ("S" for "sorites" and "D" for "disjunct," since I will be examining the disjunction of the five claims in what follows).

¹ Instead of disjunctions we could use material conditionals (e.g. "If anyone worth less than \$1 is poor, then anyone worth less than \$2 is poor"). I stick with disjunctions because many philosophers seem to be wary of material conditionals and I want to avoid those discussions.

SD₁: \sim SC₁

SD₂: \sim SC_{LAST}

SD₃: \sim [SC₂ & SC₃ & ... & SC_{LAST-1}]

SD₄: SC₁ & SC₂ & ... & SC_{LAST} & there is no truth-preserving derivation to a contradiction

SD₅: SC₁ & SC₂ & ... & SC_{LAST} & there is a truth-preserving derivation to a contradiction

The reader can verify with mere sentential logic that the disjunction of the five SDs is true. Thus, *the disjunction of the five SDs is logically true*. (More on this alleged fact in Section 9.) This is not the doxastically distressing disjunction I referred to earlier in this section. We encounter that disjunction in Section 5.

Things are going to get complicated in what follows, so it's appropriate to give a sense of the road before us. There are the "sorites claims," the SCs. Then there are the "sorites disjuncts," the SDs. I will be focusing on the latter. I will argue that each SD is philosophically counterintuitive. You might think the conclusion of those arguments is simple: Since the disjunction of the five SDs is true (as pointed out in this section), and each is philosophically counterintuitive (Section 4), this proves that some philosophically counterintuitive claim is true (*viz.* one of the SDs). Unfortunately, that's not right; there are some linguistic complications that must be addressed (Section 4). But in Section 5 we will construct a "doxastically distressing disjunction" that is obviously true (given the arguments in Sections 3 and 4) and yet it's also obvious that each disjunct is philosophically counterintuitive; hence, we have a *philosophical* argument that is *knockdown* (in senses to be discussed in Section 9) and concludes that some *violently counterintuitive* claim is true. Section 6 will consider an objection, one that doesn't require any modification of the doxastically distressing disjunction. Sections 7 and 8 take the proof method that I used on the sorites and apply it to a couple of other paradoxes in order to generate a couple more doxastically distressing disjunctions. The rest of the Element formulates and defends my ten theses.

4 Each Disjunct Is Philosophically Counterintuitive

In this section, I examine only the *obvious, immediate* philosophically interesting consequences of each disjunct, SD₁–SD₅. I will not be arguing that any particular disjunct is true (or not true). Instead, my commentaries on the disjuncts have two purposes:

- articulate the obvious, immediate philosophically interesting consequences of the disjuncts
- prove that each disjunct is *philosophically counterintuitive*, pretty much no matter how one reasonably makes that notion precise so that it comes out useful.

One can precisify “philosophically counterintuitive” as follows: A claim is philosophically counterintuitive at a time if and only if (i) if the claim is true, then a great many of our ordinary commonsensical beliefs and/or a significant portion of our most confidently held ordinary beliefs or belief-dispositions are false, or (ii) if the claim is true, then key philosophical ideas held by a large portion of philosophers at that time are false. Under (i), we have claims such as “Mom knows you bought a motorcycle,” “Kat believes global warming isn’t happening,” and “There are eight trees along the backyard property line.” Under (ii), we have claims such as “Modus ponens is truth preserving,” “No contradictions are true,” “The fact that ‘Bertrand Russell’ refers is not a brute fact.” Some claims might fall into both categories; it won’t matter. I don’t know if “philosophically counterintuitive” is the best term. “Philosophically significant” and “philosophically consequential” were considered. Set aside aptness of vocabulary. I will comment on this notion further in Section 12, addressing its philosophical significance (e.g. why should we care if a truth is “philosophically counterintuitive”?).

SD₁: It’s not the case that anyone with less than \$1 is poor.

It might seem fairly obvious that SD₁ is philosophically counterintuitive. It is saying that some people with virtually no money aren’t poor. That’s about as counterintuitive as a claim can get. It might not be as counterintuitive as “2 + 2 = 576” or “I am not conscious at all, in any sense whatsoever,” but it’s still highly counterintuitive. If it’s not the case that anyone (today) with less than \$1 (US) is (financially) poor, then a great many of our ordinary commonsensical beliefs, or a significant portion of our most confidently held ordinary beliefs or belief-dispositions, are not true. Thus, SD₁ is philosophically counterintuitive.

There is, however, a way that this argument might be unsound. Suppose SD₁ is true, so “Anyone with less than \$1 is poor” isn’t true as that sentence is used in this work of philosophy. Even so, perhaps that sentence is true when used in ordinary discourse. And if it is true in ordinary discourse, then its being false in philosophical discourse may not be nearly as counterintuitive. Just because SD₁ is false in this Element won’t mean, necessarily, that a great many of our ordinary commonsensical beliefs, or a significant portion of our most confidently held ordinary beliefs or belief-dispositions, are not true.

This type of situation happens. For instance, ordinary discourse employing “miracle,” “conscious,” “believe,” and “justified” might not match up with philosophical discourse using the same terms, since the two discourses often (not always) employ different relevant meanings (i.e. ones that change truth-value even in extensional sentential contexts). The following sentences, appearing in philosophical discourse, should strike a philosopher as having a real chance at being false:

If her belief was not justified, then “Her belief was justified” isn’t true in ordinary discourse.

If there are no miracles, then “There are miracles” isn’t true in ordinary discourse.

When I argued that SD_1 is philosophically counterintuitive, it’s arguable that I tacitly employed a similar premise in my argument:

1. If SD_1 is true, then “It’s not the case that anyone with less than \$1 is poor” is true in the discourse I am using right now in this Element.
2. If that sentence is true in the discourse I’m using right now in this Element, then it’s true in ordinary discourse.
3. If it’s true in ordinary discourse, then a great many of our ordinary commonsensical beliefs, or a significant portion of our most confidently held ordinary beliefs or belief-dispositions, are not true.
4. Hence, by (1)–(3), if SD_1 is true, then a great many of our ordinary commonsensical beliefs, or a significant portion of our most confidently held ordinary beliefs or belief-dispositions, are not true.

Sure enough, if premise (2) is true, then SD_1 is philosophically counterintuitive (since (1) and (3) are true and $\{(1)–(3)\}$ entails (4), as the conditionals (1)–(3) are material). But is (2) true?

I suppose that if one is sentimental enough, and in the sentimentality has a certain direction so to speak, one might be tempted by the idea that philosophical discourse is quite different from ordinary discourse: more sophisticated, more elegant, urbane, almost divine, ontologically and explanatorily fit, and extraordinarily charming, good-looking, and sexually impressive. Or perhaps it goes in the other direction: philosophical discourse is ill-suited for ever finding truth, hopelessly ambiguous, ugly, smelly, and crude. Joking aside, though, there are five excellent reasons for thinking that (2) is true.

First, there aren’t any problematic terms here similar to “miracle” or “justified”; hence, that particular reason to be suspicious of (2) fails to apply. Second, I am explicitly saying – right now, if you like – that I am using, in this Element, ordinary English – unless, of course, I *supplement* it with philosophical jargon,