

# 7. Why Philosophy Can Overturn Common Sense<sup>1</sup>

*Susanna Rinard*

## INTRODUCTION

Many philosophers have a rather limited view of what our discipline could hope to achieve. They think that philosophical arguments cannot rationally overturn our pre-theoretical common sense convictions. Here, for example, is Kit Fine:

In this age of post-Moorean modesty, many of us are inclined to doubt that philosophy is in possession of arguments that might genuinely serve to undermine what we ordinarily believe.<sup>2</sup>

And David Lewis:

One comes to philosophy already endowed with a stock of opinions. It is not the business of philosophy either to undermine or justify these preexisting opinions to any great extent, but only to try to discover ways of expanding them into an orderly system.<sup>3</sup>

Other advocates of positions of this kind include Lycan 2001, Kelly 2005, and Kelly 2008. On the other hand, some philosophers take the opposing view. They present and endorse philosophical arguments against various claims of common sense. For example, Unger 1975 argues that no one knows anything at all. Van Inwagen 1990, Dorr 2002, and others argue that ordinary objects like tables and chairs do not exist.

Finally, I'll include some remarks on this topic by Kant and Hegel:

<sup>1</sup> Thanks to Alex Byrne, Andrew Graham, David Gray, Daniel Greco, Bradford Skow, Agustin Rayo, Robert Stalnaker, John Stamm, three anonymous referees for Oxford Studies in Epistemology, and audiences at the MIT MATTI Reading Group and the MIT Epistemology Reading Group. Thanks especially to Miriam Schoenfeld for our many helpful conversations on these topics. Thanks to Andrew Graham for providing the Kant and Hegel quotes. Special thanks to Roger White—this paper has benefited greatly from his many suggestions and comments. Finally, I would also like to thank Thomas Kelly for writing two detailed and fascinating papers on this topic, the reading of which inspired me to write this paper. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship.

<sup>2</sup> Fine (2001, 2).

<sup>3</sup> Lewis (1973, 88).

To appeal to ordinary common sense... is one of the subtle discoveries of recent times, whereby the dullest windbag can confidently take on the most profound thinker and hold his own with him.<sup>4</sup>

Since the man of common sense makes his appeal to feeling, to an oracle within his breast, he is finished and done with anyone who does not agree; he has only to explain that he has nothing more to say to anyone who does not find and feel the same in himself. In other words, he tramples underfoot the roots of humanity.<sup>5</sup>

As for myself, I wouldn't go so far as to say that Lewis, Fine, and company are dull windbags. And I am not yet convinced that they have trampled underfoot the roots of humanity. However, my view shares some of the spirit of these remarks. For I *do* believe, contra Lewis, Fine, etc., that philosophy *can* overturn common sense. It is the aim of this paper to defend that position. (I will not begin by defining "common sense." If this concerns you, please see this footnote.<sup>6</sup>)

The paper has two distinct and easily separable parts. In part one, I present and endorse a positive argument for the claim that philosophy can overturn common sense. My opponents and I agree that science can overturn common sense. But, I claim, every scientific argument relies on assumptions that are highly theoretical, even philosophical. If a scientific argument against a common sense proposition is to succeed, then its philosophical assumptions must be more worthy of belief than the common sense proposition under attack. But this means that there could be a philosophical argument against common sense, each of whose premises is just as powerful, epistemically, as the scientist's philosophical assumptions. If the scientific argument can succeed, then so, too, can the purely philosophical argument, and so philosophy *is* capable of overturning common sense.<sup>7</sup>

<sup>4</sup> Kant (2008, 9).

<sup>5</sup> Hegel (1977, 43).

<sup>6</sup> I will not take it upon myself here to provide any general definition of "common sense," and I don't think this is required for my arguments to go through. My opponents are committed to the view that there is an important distinction between what is common sense and what isn't, since they believe that this distinction plays an important epistemic role. Since I deny the latter, I am not committed to the former. That said, in the course of arguing against my opponents, there are a few places in which my arguments do rely on a claim about what counts as common sense. In each such case, I think that most if not all of my opponents would agree. For example, in the second half of the paper, I assume that certain propositions about the external world, like "I have hands," count as common sense. These are taken to be paradigm examples of common-sense propositions by all parties to the debate. Also, in the first half of the paper, I assume that special relativity conflicts with common sense. Although perhaps slightly more contentious, this is also generally accepted. Moreover, see section 3, response to objection 5, for a response to the objection that special relativity does not conflict with common sense.

<sup>7</sup> One might wonder whether this conclusion by itself is enough to seriously challenge a broadly Moorean outlook. Shouldn't the reasonable Moorean admit that there might possibly be one or two cases in which philosophy can overturn common sense? Surely we should interpret the Moorean as claiming at most that, in general, philosophy can't overturn common sense. I'd like to direct those worried about this issue to the last three paragraphs of section 2. One central aspect of the Moorean outlook, however exactly it is interpreted, is that we are supposed to be

In part two, I undermine some of the main reasons philosophers have given for the opposing view that philosophy can't overturn common sense. First I consider the Moorean idea that common sense propositions are more plausible than philosophical claims. Then I turn to the view, defended in Kelly 2005, that one should retain belief in judgments about particular cases at the expense of general principles that conflict with them. Finally I consider a version of reflective equilibrium, defended in Harman 2003, in which conservatism plays an important role. In each case, I argue that either the view in question is false, or that it fails to provide an independent motivation for the claim that philosophy can't overturn common sense.

The aim of this paper, then, is to argue for and defend the claim that philosophy can overturn common sense. If I am right, then I think this infuses the project of philosophy with new importance and urgency. Almost everything we think, say, and do presupposes some proposition of common sense. If the business of philosophy is, at least in part, to inquire into the truth of these propositions—with the real possibility left open that we may find good reasons for rejecting them—then philosophy is highly relevant to almost every aspect of our daily lives.

#### 1. INTRODUCTION TO PART ONE: PHILOSOPHY CAN OVERTURN COMMON SENSE

In this first part of the paper I will provide a positive argument for the claim that philosophy can overturn common sense. In its simplest form, the argument is as follows:

- (1) Science can overturn common sense.
- (2) If science can overturn common sense, then so can philosophy.
- (3) Therefore, philosophy can overturn common sense.

This argument is not original to me. Indeed, it is considered, and rejected, by many of my opponents. My main contribution will come in the form of my particular defense of premise (2), for it is this premise that is generally rejected by the advocates of common sense. Premise (1) is widely accepted, and the argument is valid, so I will focus primarily on defending premise (2) (though see section 3, reply to objection 5 for a defense of premise (1)).

It will be helpful to begin by considering my opponents' argument against premise (2). Here are some relevant quotes, starting with one from William Lycan:

Common sense beliefs can be corrected, even trashed entirely, by careful empirical investigation and scientific theorizing...No *purely philosophical* premise can ever

able to dismiss philosophical arguments for skepticism on the grounds that they conflict with common sense. I claim at the end of section 2 that, if my arguments are accepted, we can't do this. So I think that my arguments really do take the wind out of the Moorean's sails. My reply to objection 5 in section 3 is also relevant here.

(legitimately) have as strong a claim to our allegiance as can a humble common sense proposition such as Moore's autobiographical ones. Science can correct common sense; metaphysics and philosophical "intuition" can only throw spitballs.<sup>8</sup>

Anil Gupta expresses a similar sentiment:

Any theory that would wage war against common sense had better come loaded with some powerful ammunition. Philosophy is incapable of providing such ammunition. Empirical sciences are a better source.<sup>9</sup>

The idea here—which is also found in Kelly 2008—can be summarized as follows: science, unlike philosophy, can appeal to empirical, observational evidence. When science undermines common sense, it does so by appealing to direct observation. When the philosopher attempts to undermine common sense, however, she can appeal only to highly theoretical premises, which are less powerful, epistemically, than observational evidence. So scientific arguments against common sense are more powerful than philosophical argument against common sense.

There are many concerns one might have about this argument. First, it is notoriously difficult to distinguish observational and theoretical claims. Second, even supposing one can do so, it is not clear that observational claims really are epistemically stronger than theoretical ones.

However, I want to set these worries aside, and focus on what I think is a deeper flaw in this argument. I agree with my opponents that scientific arguments, unlike philosophical arguments, appeal to observational claims. However, I will argue that scientific arguments must also rely on highly theoretical assumptions that are just as far removed from observation as the kinds of claims typically appealed to in philosophical arguments against common sense. Indeed, many of these theoretical scientific assumptions are straightforward examples of typical philosophical claims. An argument is only as strong as its weakest premise. So if a scientific argument is to succeed in undermining common sense, then each of its premises, individually, must be more epistemically powerful than the common sense proposition it targets.<sup>10</sup> Since, as I claim, the scientific argument relies crucially on a philosophical assumption, this philosophical assumption must be more powerful than the common sense proposition. But if one philosophical claim can be more powerful than a common sense proposition, then there could be an argument consisting entirely of philosophical claims, each of which is more powerful than the common sense proposition whose negation they entail. If so, then philosophy can overturn common sense.

The argument I just sketched appeals to the following claim:

<sup>8</sup> Lycan 2001.

<sup>9</sup> Gupta (2006, 178).

<sup>10</sup> A is more epistemically powerful than B just in case, if forced to choose between A and B, one should retain belief in A and give up belief in B.

*Science Requires Philosophy (SRP)*: Scientific arguments against common sense rely crucially on philosophical assumptions.

The next section will be devoted to defending this claim.

## 2. SCIENCE REQUIRES PHILOSOPHY (SRP)

I will begin my case for SRP by considering what is perhaps the most widely accepted example of a case in which science has succeeded in overturning common sense: special relativity. Most philosophers—including many who think that philosophy can't overturn common sense—believe that special relativity is true, and agree that special relativity conflicts with common sense. So if I can show that the scientific argument for special relativity relies crucially on a philosophical assumption, then this, in combination with the fact that an argument is only as strong as its weakest premise, will suffice to show that philosophy can overturn common sense (as explained in more detail in the previous section).

We needn't get too far into the technical details, but some basic knowledge of the case will be helpful. Consider a simultaneity proposition like this one: Joe's piano recital and Sarah's baseball game were happening at the same time. Pre-theoretically, we would think that a proposition like this one, if true, is objectively and absolutely true; its truth is not relative to a particular person or thing. "Licorice ice cream is delicious" may be true for Joe but not for Sarah, but the simultaneity proposition, we would normally think, is true absolutely if true at all. However, according to special relativity (SR), this is not the case. SR says that there are many different reference frames—each object has its own—and the very same simultaneity proposition may be true in one reference frame but not another. Moreover, there's no particular reference frame that has got it right. Each reference frame has the same status; none is more legitimate than the others. So, special relativity conflicts with the common sense idea that simultaneity claims are absolute.

Now, there is an alternative scientific hypothesis—the so-called neo-Lorentzian view—that is empirically equivalent to special relativity but which does not conflict with the common sense idea that simultaneity is absolute. Special relativity and neo-Lorentzianism agree on almost everything. In particular, they agree on all observational propositions—there is no possible experiment that could decide between them. The main difference is that according to neo-Lorentzianism, one of the reference frames is privileged in the sense that it gets the simultaneity facts right. On this view, one particular reference frame is objectively and absolutely correct. So the neo-Lorentzian view vindicates the common sense idea that simultaneity is absolute.

Most scientists—and most philosophers—believe that special relativity, rather than the neo-Lorentzian view, is true. They say that the neo-Lorentzian view is unnecessarily complex. It posits an additional metaphysical fact—a fact about which reference frame gets things absolutely correct—that doesn't

make a difference to the empirical predictions of the theory. Special relativity “gets the job done” with less machinery. So if we follow Ockham in thinking that simpler hypotheses should be given more credence than complex ones, we will give more credence to special relativity than to the neo-Lorentzian view. This, in fact, is what most scientists and philosophers do, and for this reason they give up the common sense idea that simultaneity is absolute.

With these facts on the table, we can now ask the crucial question: does the argument from special relativity against the absoluteness of simultaneity rely crucially on a philosophical assumption? I think that it does. In particular, it relies on the philosophical assumption that simpler hypotheses should be preferred over complex ones. Anyone who gives up the view that simultaneity is absolute on the basis of special relativity must have a reason for preferring special relativity to the neo-Lorentzian view. The reason standardly given is that special relativity is simpler. Without the further claim that simpler theories should be preferred, we simply don’t have any reason to give up the common sense idea that simultaneity is absolute. So the defender of special relativity must think that a philosophical assumption—the claim that simpler theories should be preferred over complex ones—is more epistemically powerful than a common sense proposition.

Here’s another way to make the point. Suppose that this philosophical assumption were *not* more powerful than the common sense claim. In that case, one should reason as follows: well, the idea that simpler theories are preferable does have some plausibility to it. But if this is true, then we have to prefer special relativity to the neo-Lorentzian view, since it is simpler. But this would force us to give up the common sense idea that simultaneity is absolute. This common sense claim is more powerful than the philosophical assumption that simpler theories are preferable. So, I will retain belief in the common sense claim, give up my preference for simpler theories, and then believe what the empirical evidence forces me to believe, namely, that the neo-Lorentzian view is true. Most philosophers, however, do not reason in this manner. Since they think we should accept special relativity, I conclude that they must think that the philosophical preference for simplicity is more powerful than the common sense notion that simultaneity is absolute.

So, we see that the insistence by Kelly 2008, Lycan 2001, and Gupta 2006 that observational evidence is more powerful than philosophical claims, and their pointing out that science, unlike philosophy, appeals to observational evidence, is beside the point. No matter how much observational evidence is appealed to in a scientific argument against common sense, as long as the argument relies crucially on a philosophical assumption, then, if the argument is to succeed, this philosophical assumption must be more powerful than the targeted proposition of common sense. If so, then there could be a successful argument against common sense that relies only on philosophical assumptions; and if so, then philosophy is capable of overturning common sense.

The next section of this paper will be devoted to replying to a variety of objections to the argument I’ve just given. But first, before closing this section,

I'll briefly address the question of how general this phenomenon might be. That is, is the argument for special relativity unique among scientific arguments in its reliance on philosophical assumptions? Or is the phenomenon more widespread?

I think there is a general reason to think that the phenomenon is widespread. Scientific arguments against common sense typically proceed by noting that a currently accepted scientific hypothesis is in conflict with common sense. However, scientific hypotheses are generally not logically entailed by the data that support them. Moreover, it is usually only the hypothesis as a whole that conflicts with common sense, rather than the data themselves. This is true in many other commonly cited examples of the overturning of common sense by science: astronomical hypotheses according to which the Earth is round, rather than flat, and according to which the Earth orbits the Sun, rather than vice versa; and the hypothesis that tables, chairs, and other objects are mostly empty space, rather than solid.

Since it is only the hypothesis as a whole and not the empirical data by themselves that conflicts with common sense, there will always exist an empirically equivalent competitor hypothesis that vindicates common sense. If so, then a philosophical assumption will be required if the non-common-sensical theory is preferred. Such an assumption will likely be an epistemological principle about theory choice, such as the claim that one hypothesis explains the data better and should for that reason be preferred; or that one hypothesis is simpler, more internally coherent, or better unified with other accepted theories, and that these constitute reasons for preferring it; etc. So, I think that the reliance of science on philosophy is not an isolated phenomenon restricted to a few cases like special relativity, but is rather the norm.

Let's take stock. I have argued that the paradigm example of a successful scientific argument against common sense—the argument for special relativity—relies crucially on a philosophical assumption, namely the assumption that simpler hypotheses should be preferred over complex ones. Anyone who accepts special relativity on the basis of the scientific argument for it is committed to thinking that this philosophical assumption is more epistemically powerful than the common sense idea that simultaneity is absolute. If so, then there could be a successful argument against a common sense proposition that relied only on philosophical assumptions. If each of its premises is at least as powerful as the claim that simpler theories should be preferred, then the philosophical argument against common sense will succeed.

One upshot of this is that we can't dismiss arguments like the argument for external world skepticism just on the grounds that its premises are purely philosophical. Rather, we must carefully consider the status of each of its premises in comparison to the common sense claim that we know we have hands—and, crucially, in comparison to the status of the philosophical assumptions required by the scientific arguments that one accepts. If the premises of the skeptical argument are as powerful as, for example, the claim

that simpler theories should be preferred, then the skeptical argument will succeed.

It is not my purpose here to deliver a final verdict on the success of the skeptical argument,<sup>11</sup> so I will not undertake an in-depth comparison here. However, I will note that, on the face of it, things don't look at all bad for the skeptic. Take, for example, one of the key premises in one version of the argument for skepticism: the claim that propositions about the way things appear to us—for example, the proposition that it appears as though I have a hand—are evidentially neutral between the hypothesis that things really are the way they appear (I really do have a hand) and the hypothesis that I am a brain in a vat being fed misleading appearances as of a hand. This claim is extremely compelling. How could the appearance of a hand be evidence for one of these hypotheses over the other, when both predict that I would have exactly this appearance? Compare this claim, now, with the claim that we ought to prefer simpler theories over complex ones. While this claim is accepted by many philosophers, it seems to me if anything *less* obviously correct than the skeptic's premise just mentioned. If the preference for simplicity is powerful enough to overturn common sense, then it seems to me that the skeptic's claim is as well.

Of course, there is much more that could be said on this topic—for example, one might try to argue in response that the common sense claim that simultaneity is absolute was antecedently less powerful than the claim that I know I have hands, and so it may not suffice for the skeptic's premises to be as powerful as the scientist's philosophical assumptions. My point here is just that once we have seen that, as in the case of special relativity, philosophical assumptions *can* be more powerful than common sense, skeptical arguments and other philosophical attacks on common sense can no longer be dismissed out of hand. A careful and serious investigation into the epistemic status of their premises needs to be undertaken, and, at the outset, it is not at all clear that these premises won't in the end prove powerful enough to overturn common sense.

### 3. OBJECTIONS AND REPLIES

**Objection 1:** Your argument (says the objector) presupposes that the scientific argument for special relativity relies crucially on the philosophical assumption that simpler theories should be preferred to complex ones. However (says the objector), not all arguments for special relativity rely on this assumption. The following is a perfectly valid argument: (1) [empirical scientific data];

<sup>11</sup> I do so in Rinard (ms). I argue that it is not rational to accept the argument for external world skepticism because anyone who does so is ultimately committed to accepting skepticism about complex reasoning on the basis of a complex argument (via skepticism about the past), a position that is self-undermining.



(2) If [empirical scientific data], then special relativity is true; therefore, (3) special relativity is true.

*Reply:* Let's think more about the status of premise (2) in the objector's argument. We'll assume that the person considering the argument is a scientist or philosopher aware of the existence of the neo-Lorentzian view. What reason could such a person have for believing (2)? After all, the empirical data are entailed by both special relativity and neo-Lorentzianism. It seems to me that such a person must think that special relativity has some other feature that neo-Lorentzianism does not have, such that hypotheses with that feature should be given greater credence. But, whatever the feature, the claim that hypotheses with this feature should be given greater credence will be a philosophical, epistemological claim. If one doesn't believe a philosophical claim of this kind, then one would not believe (2), and the argument wouldn't go through. So I claim that any argument for special relativity must rely at least implicitly on a philosophical assumption.

*Objection 2:* I concede (says the objector) that the scientific argument for special relativity relies on a philosophical assumption. However, we have more reason to believe the philosophical assumptions that are typically appealed to in scientific arguments against common sense than the philosophical assumptions that typically appear in philosophical arguments against common sense. This is because science has such an impressive track record. Every time we use a laptop or walk over a bridge, we are getting evidence that scientists know what they are doing, and that we should believe whatever theories and philosophical assumptions are endorsed by science. We have no similar reason for believing the assumptions made by philosophers, since philosophy as a discipline does not have a similar track-record.

*Reply:* According to the objector, it is in virtue of science's long and impressive track-record of success that we should believe the philosophical assumptions that appear in scientific arguments like the argument for special relativity. If the objector is right, this means that in the absence of such a track record, we would have no reason to believe these assumptions. But I don't think this is right, and I don't think my opponents would agree with this either.

Consider a hypothetical scenario in which human history went quite differently. Let's suppose that the very first scientist managed to acquire the empirical evidence that actual scientists now take to constitute their reason for believing special relativity. Let's suppose that after reflecting on this evidence, the first scientist developed the theory of special relativity. She also developed and considered the neo-Lorentzian view, but reasoned that special relativity should be preferred because it was simpler and more elegant. Now, given that this scientist is working in a society that does not have a long track-record of the success of science, according to the objection just given, this scientist has no reason to believe her philosophical assumption that simpler theories are to be preferred. However, I think the scientists and philosophers who believe

special relativity today would agree that this first scientist would be entirely justified in assuming that simpler theories should be preferred, and giving up her common sense beliefs on that basis. If so, then a long track-record is not required for science to overturn common sense, and the objection fails.

**Objection 3:** You (says the objector) have argued that anyone who gives up the common sense idea that simultaneity is absolute must do so on the basis of an argument that relies crucially on philosophical assumptions. But you assumed that the person in question was a scientist or philosopher familiar with the neo-Lorentzian view. However, many lay people have given up the absoluteness of simultaneity without having the faintest idea of what neo-Lorentzianism is, or even what special relativity is, just on the basis of the testimony of scientists or teachers. Surely they didn't have to rely on any assumptions about the relative merits of simple and complex scientific theories, but it was rational nonetheless for them to give up the common sense idea that simultaneity is absolute.

**Reply:** First, note that my argument does not require that *everyone* who rationally gives up the absoluteness of simultaneity must rely on philosophical assumptions. It is enough that one could rationally do so on the basis of an argument that does rely on philosophical assumptions, as that alone is sufficient to show that philosophical assumptions can be more powerful than common sense.

However, as a matter of fact, I do think the layperson described by the objector is relying, at least implicitly, on some philosophical assumptions, although they may be quite different from the assumptions relied on by the scientist. For example, she is relying on the assumption that testimony is a legitimate way to acquire justified beliefs. This is an epistemological claim.

**Objection 4:** The scientists' philosophical assumptions are more powerful than the philosophers' because there is more agreement about them.

**Reply:** As in my reply to objection 2, we can consider a hypothetical scenario in which there is no established scientific community, nor even an established intellectual community. According to the objector, a philosophical assumption is strong enough to overturn common sense only when there is consensus about it. But I think even my opponents would agree that a lone scientist, in ignorance of what anyone else thinks of the idea that simpler theories are preferable, could, if in possession of the right empirical evidence, rationally come to believe special relativity by relying in part on philosophical claims. This shows that it is not in virtue of the consensus in the scientific community that philosophical assumptions can be more powerful than common sense.

At this point, the objector may concede that consensus is not required for one to be justified in the assumption that simpler theories should be preferred to complex ones. However, the objector may then claim that the presence of significant disagreement would suffice to undermine one's justification in this philosophical assumption. Moreover, says the objector, since

there *is* significant disagreement about the philosophical assumptions that are employed in philosophical arguments against common sense, this undermines any justification we might otherwise have had for accepting these arguments. On this view, philosophy can't overturn common sense (even though science can) because there is significant disagreement about the philosophical assumptions made in philosophical arguments.

Once again, however, I think the objector is committed to some implausible claims about certain hypothetical scenarios. Suppose that, when special relativity was initially presented to the scientific community, there was significant disagreement about whether the theory should be accepted. Many—perhaps most—scientists thought the theory too absurd to take seriously, and so did not give up their belief that simultaneity is absolute. The objector is committed to thinking that, in such a case, the initial proponent of special relativity would not be justified in believing it. However, this does not seem right. It *can* be rational for scientists (and philosophers) to maintain their views even in the face of significant disagreement.<sup>12</sup> Since disagreement would not be sufficient to undermine justification in the philosophical assumptions required by scientific arguments against common sense, it is not sufficient to undermine justification in the premises of philosophical arguments against common sense.

**Objection 5:** I concede (says the objector) that special relativity has overturned our belief that simultaneity is absolute. The claim that simultaneity is absolute may be plausible; however, I think it is not as basic, as robust, or as central to our way of thinking as the kind of common sense beliefs that philosophers attempt to undermine, such as the belief that I know I have hands or the belief that tables exist. Science may be able to undermine widely accepted propositions that are highly plausible—like the proposition that tables are solid, that the Earth is flat, that the Sun orbits the Earth, and that simultaneity is absolute—but even science couldn't undermine the real “hard core” of common sense that philosophical arguments target. So examples from science like the example of special relativity don't really go any way towards making

<sup>12</sup> If we also assume that those with whom one disagrees are epistemic peers in the relevant sense, then some epistemologists would disagree with this claim; see, for example, Elga (2007) and Christensen (2007). To someone with that kind of view—someone who denies that it can be rational for a scientist to believe a proposition in the face of significant peer disagreement—I offer the following, alternative response to this objection. According to this objector, it is only because of the widespread disagreement about certain philosophical claims that they are not strong enough to overturn common sense. Thus, according to the objector, it's entirely possible that a philosopher working in isolation could rationally give up common sense beliefs on the basis of philosophical arguments. If so, then the reason why, according to the objector, philosophy can't *actually* undermine common sense is just due to an accidental, contingent fact about the existence of disagreement in the field. But this does not seem to be in the spirit of their view. I would think that many of my Moorean opponents would want to insist that even the lone philosopher should not give up common sense beliefs on the basis of philosophical arguments. If so, then they would not want to endorse objection 4.

it more plausible that philosophical arguments like the skeptical argument could succeed.

*Reply:* This objector is objecting to premise (1) of the simple statement of my argument as it appears at the beginning of section 1, which is the premise that science can overturn common sense. According to the objector, there is a “hard core” of common sense that can’t be overturned by any sort of inquiry, either scientific or philosophical, and this “hard core” includes propositions in conflict with external world skepticism and ontological nihilism.

Many of my opponents accept premise (1), and so I have simply presupposed it up until this point, for reasons of dialectical effectiveness. However, I think this objector makes a good point, and so I’ll take up the issue here.

I don’t want to rest my case on this, but I want to begin by saying that it’s not obvious that special relativity isn’t in conflict with propositions just as central and basic as the proposition that I know I have hands. Consider an ordinary simultaneity proposition like the following: Andrew and I are brushing our teeth at the same time. One might think—indeed, this is my view—that this proposition, as ordinarily understood, could be true only if it is objectively and absolutely true that Andrew and I are brushing our teeth at the same time.<sup>13</sup> If so, then it is not merely the abstract and theoretical-sounding proposition that simultaneity is absolute that is in conflict with special relativity; rather, this theory is in conflict with propositions as ordinary, simple, and plausible as any day-to-day simultaneity claim.<sup>14</sup>

Consider, also, the astronomical hypothesis that the Earth orbits the Sun. One might be inclined to think that if this hypothesis is true, then my bed is not in the same place as it was yesterday. After all, the Earth has moved somewhat, and so the region of space formerly occupied by my bed is now occupied by something else. But the claim that my bed is in the same place it was yesterday seems to me on par with the claim that I know I have hands. This gives us a reason for being skeptical of the objector’s claim that the propositions overturned by science are not part of the “hard core” of common sense that philosophers have attempted to undermine.

But I don’t want to rest my case on these claims. Rather, I will argue that there could be successful empirical arguments against the very claims that, according to the objector, are in the “hard core” of common sense, and that these arguments rely crucially on philosophical assumptions.

<sup>13</sup> Similarly, although I won’t attempt to argue for this here, one might think that ethical relativism is off the table because, as ordinarily understood, propositions like “Torture is wrong” are true only if true absolutely. On this view, part of what one is asserting when one asserts that torture is wrong is that torture is wrong according to all legitimate perspectives.

<sup>14</sup> Special relativity also entails that nothing can go faster than the speed of light. This may also seem to conflict with some very basic common-sensical ideas. For example, suppose I am on a train going almost as fast as the speed of light (but not quite) and I shoot a gun in the direction of the train’s movement whose bullets go half as fast as the speed of light. It seems like the bullet should end up going faster than the speed of light. But according to special relativity, this is impossible. (Thanks to Bradford Skow for suggesting that I include this example.)

Consider, for example, one of the common sense claims that the skeptic aims to undermine: I know I have hands. Epistemologists agree that one could get empirical evidence against this claim. For example, suppose one is told by a reliable source that doctors have found a way to cure cancer by manufacturing a drug that requires some part of a real human hand. People are being asked to donate their hands to the cause, so that enough of this drug can be manufactured to cure everyone of cancer. Those who agree to donate their hands are told they will undergo a surgical procedure under general anesthesia which involves the removal of their hands and the replacement of them with fake hands, which look and feel exactly like real hands, and from which there is no recovery period. (We can imagine that surgery has become quite advanced.) You agree to donate your hands. When you wake up in the hospital bed, you look down at the ends of your arms, and find that what appear to be your hands look and function exactly as they always have, just as you were told they would. In this case, I think you should believe that you do not have real hands (just fake hands), and so give up the common sense belief that you know you have hands. Moreover, in giving up this belief you are (at least implicitly) relying on the epistemological assumption that the fake-hand hypothesis is better supported by your evidence than the empirically equivalent conspiracy theory according to which the doctors are all lying to you and your hands were never removed in the first place.

Here's another example in which one could get empirical evidence against the common sense belief that one has hands. Suppose you wake up one morning and find a ticker-tape across the bottom of your visual field. The tape reads: "Your brain has been removed from your body and put in a nutrient-filled vat. Your sensory experiences are being fed to you by our vat-tending scientists." The tape goes on to make all sorts of predictions about what your sensory experience will be like, and these predictions all come true.<sup>15</sup> In such a case, I think one should believe that one is a brain in a vat, and so give up all of the common sense beliefs that are incompatible with that claim. But, once again, in doing so one must rely at least implicitly on an epistemological claim according to which the brain in a vat (BIV) hypothesis is more worthy of belief, given your evidence, than the hypothesis that things really are as they seem, and that the ticker-tape is not reliable. (Perhaps, according to this alternative hypothesis, you are hallucinating the ticker-tape due to some kind of psychological ailment.)

It is worth pointing out that the kind of epistemological assumptions featured in these cases are very similar in kind to the types of epistemological assumptions typically appealed to in philosophical arguments for skepticism. For example, the skeptic may employ the premise that your current sensory evidence is neutral between the BIV hypothesis and the hypothesis that things

<sup>15</sup> This example is not original to me, but unfortunately I can't remember where I first heard it.

are as they appear. The epistemological assumptions appealed to in the above-described empirical arguments against common sense are claims of a similar kind: claims about which hypotheses about the nature of the real world are best supported by the evidence of your senses.

This completes my reply to this objection. I have argued that there could be successful empirical arguments against the very same common sense propositions that philosophical arguments seek to undermine. Moreover, these empirical arguments rely crucially on philosophical assumptions. So, in the relevant sense, science can overturn common sense and premise (1) of the argument remains intact.

This concludes part one of the paper. The overall goal of this first part was to argue that philosophy can overturn common sense, since science can. Scientific arguments against common sense, I claim, rely crucially on philosophical assumptions. For example, the argument for special relativity relies crucially on the assumption that simpler hypotheses should be preferred over complex ones. Since an argument is only as strong as its weakest premise, if these scientific arguments succeed, it must be the case that the philosophical assumptions on which they rely are stronger than the common sense propositions they target. But if so, then there could be an argument against common sense whose premises consist entirely of philosophical claims, each of which is stronger than the targeted common sense proposition. So philosophy can overturn common sense, since science can.

#### 4. INTRODUCTION TO PART TWO

In this second half of the paper I will consider some of the most common motivations for my opponents' view that philosophy cannot overturn common sense. Each motivation relies crucially on some principle of philosophical methodology. I will consider the following three methodological principles: (1) common sense propositions enjoy greater plausibility than philosophical claims, and thus should be given priority when they conflict (defended most famously by Moore 1962a, 1962b, 1962c); (2) general principles should be rejected when they conflict with a large number of judgments about particular cases (defended in Kelly 2005 as a kind of Mooreanism and also, I think, suggested by certain remarks in Goodman 1983); (3) conservatism—minimizing change—is an important aspect of rational belief revision (defended in Harman 2003 as a version of reflective equilibrium and also suggested by certain remarks in Lewis 1999).

In each of these three cases, I will argue either that the methodological principle in question is false, or that it fails to provide a genuinely independent motivation for the idea that philosophy can't overturn common sense. For example, in section 7 I will present a novel objection to conservatism that, I argue, undermines Harman's principle of philosophical methodology. I will claim that, on one important notion of epistemic dependence, beliefs

that depend on a belief Q are completely irrelevant to whether belief in Q should be maintained. However, I will show that there is an important class of cases in which principles of methodological conservatism (like Harman's) have the consequence that whether or not belief in Q should be maintained depends in part on the nature and number of beliefs that depend on Q. I think this brings out a deep problem with the view that conservatism (minimizing overall change) is an important aspect of philosophical methodology.

Throughout the discussion of all three methodological principles, I will focus (as do the defenders of these principles) on the case of external world skepticism. The skeptic presents us with a valid argument from plausible philosophical premises for the conclusion that no one knows anything about the external world. For example, the skeptic may argue that all we have to go on in assessing propositions about the external world is the evidence of our senses, that is, propositions about how things appear to us. But, continues the skeptic, this evidence is neutral between the hypothesis that the external world really is as it seems and the hypothesis that one is a brain in a vat being deceived by a mad scientist. So, concludes the skeptic, no one knows whether they really have hands, whether there is a table before them, etc.

By presenting her argument, the skeptic has shown us that there is a contradiction between some plausible epistemological principles and our ordinary common sense beliefs about what we know. How should we revise our beliefs in light of this? Should we accept the skeptic's premises and the radical conclusion that follows from them? Or should we hold on to our pre-theoretic belief that we know many things, and reject one of the skeptic's premises? Each of the three methodological principles I will discuss yields the verdict (according to its defenders) that we should hold on to our common sense beliefs about the extent of our knowledge, and give up one of the skeptic's premises. But I will argue that these principles do not succeed in motivating this view.

##### 5. MOORE'S PLAUSIBILITY PRINCIPLE

We begin, of course, with G. E. Moore. I'll follow Lycan 2001 p. 38–39 in characterizing Moore's view roughly as follows: A common sense proposition is more plausible than the premises in a philosophical argument against it. So, if forced to choose between common sense and a philosophical premise, one should retain belief in common sense and give up the philosophical premise. Call this Moore's Plausibility Principle.

In assessing this view, I think we should first ask what is meant by "plausible." On one reading, A is more plausible than B just in case A seems pre-theoretically *prima facie* more worthy of belief than B. If this is what is meant by "plausible," then I am willing to grant that ordinary common sense may indeed be more plausible than complex abstract philosophical principles.

The trouble is that it is simply not the case that, when conflicts arise, we should always retain belief in whichever proposition seems pre-theoretically *prima facie* more worthy of belief. Sometimes one discovers by reflection that one's pre-theoretical judgments of comparative worthiness of belief were the reverse of what they should have been. (This point is also made by Conee 2001 and echoed by Kelly 2005.) For example, consider the following two propositions: (A) There are more positive integers than there are positive even integers. (B) Two sets X and Y have the same cardinality just in case there is a one-to-one mapping from the elements of X to the elements of Y. Pre-theoretically, A seems more worthy of belief than B. It just seems *obvious* that there are more positive integers than positive even integers, and B is a complicated principle that requires some thought before endorsing. So if plausibility is interpreted as suggested above, Moore's Plausibility Principle says that one should retain belief in A and reject B. However, this is exactly the opposite of what should be done. Reflection should result in a reversal of one's initial judgment of the comparative worthiness of belief of these two propositions.

For all that has been said so far, the skeptical case could be just like this one. Pre-theoretically, it may seem obvious that one knows that one has hands, and the premises of the skeptical argument may seem to be complicated claims that require thought before endorsing. However, after reflecting on the skeptical premises, they can come to seem more and more plausible until it becomes clear that they are undeniable, and the claim to knowledge must be rejected. So, although the skeptic may agree with the Moorean about the pre-theoretical, *prima facie* judgments, neither the skeptic nor the Moorean should accept the principle that these initial judgments determine which of the two propositions it is rational to give up, since reflection can sometimes reveal that one's initial judgments were the reverse of what they should have been.<sup>16</sup>

This discussion may suggest an alternative reading of the Plausibility Principle. On the first reading, which of two conflicting propositions you should give up depends on your pre-theoretic judgments about them. On the second version of the principle, which of the two propositions you should give up is not determined by your *initial* judgments; rather, it is determined by your judgments *upon careful reflection*. According to this version of the principle, one should give up whichever of the two propositions seems least worthy of belief upon careful reflection.

However, this version of the principle does not yield the result the Moorean desires. It makes the normative facts about what one should believe dependent on the contingent psychological facts about what happens to seem most

<sup>16</sup> The Monty Hall problem may provide another example of a case in which one's initial judgments of the comparative worthiness of belief are the reverse of what they should be. In general, psychology experiments like the Wason card selection task and the experiments discussed in the heuristics and biases literature (see for example Kahneman, Tversky, and Slovic 1982) provide another source of counterexamples to Moore's Plausibility Principle.



worthy of belief upon reflection. As a matter of psychological fact, upon careful reflection, each premise of the skeptical argument seems to the external world skeptic to be more worthy of belief than the common sense view that she has knowledge of the external world. So, according to this version of the plausibility principle, she should give up the common sense belief on the basis of the philosophical argument for skepticism. If so, then philosophical argument can undermine common sense, which is exactly what the Moorean seeks to deny.

We may consider a third, more normative understanding of what plausibility amounts to: A is more plausible than B just in case as a matter of fact A is more worthy of belief than B. This does not seem to help much; after all, the skeptic thinks that her premises *are* more worthy of belief than the common sense claim to knowledge, and so the skeptic will think that this version of the plausibility principle vindicates her position.

The Moorean may try to tack on the claim that the skeptic is wrong about this: that, as a matter of fact, it is the common sense claim that is most worthy of belief. Simply to insist that this is the case, however, does not provide us with an independent motivation for the Moorean view. To say that common sense propositions are more worthy of belief than philosophical claims just is to say that, in cases of conflict, we should retain belief in common sense propositions at the expense of philosophical claims, which is just to say that philosophy can't overturn common sense. What we have here is a restatement of the Moorean view, not an independent motivation for it.

I conclude, then, that each version of the plausibility principle is either false, or fails to provide a genuinely independent motivation for the view that philosophy can't overturn common sense.

## 6. GENERAL PRINCIPLES VS. PARTICULAR CASE JUDGMENTS

In his 2005 paper "Moorean Facts and Belief Revision or Can the Skeptic Win?", Thomas Kelly draws the same conclusion that I did in the previous section: Moore's plausibility principle does not successfully make the case that philosophical argumentation can't overturn common sense propositions like our ordinary claims to knowledge. However, Kelly then goes on to provide an alternative interpretation of Moore that he thinks does succeed in showing this. I think Kelly's version of Mooreanism is also unsuccessful.

Kelly begins by distinguishing two different principles of philosophical methodology: particularism and methodism.<sup>17</sup> The particularist and the methodist both begin with some initial judgments about particular cases, and

<sup>17</sup> Kelly's terminology is borrowed from Chisholm's 1973 book *The Problem of the Criterion*. Chisholm used the terms slightly differently, to mark the distinction between the epistemologist who builds his theory of knowledge around his initial views about the extent of our knowledge (the particularist) and the epistemologist who builds his theory of knowledge around his initial views about the criteria for knowledge (the methodist).

some initial judgments about general principles. The difference between the particularist and the methodist manifests itself when contradictions are discovered between the initial case judgments and general principle judgments.

In revising his beliefs in light of such contradictions, the particularist will give more weight to his initial judgments about cases than to his judgments about general principles. So, if just one of his initial case judgments is in conflict with one of his general principle judgments, he will give up the general principle and retain the case judgment. The methodist, on the other hand, will give more weight to the general principle judgment than to the case judgment. So, in a case of conflict between one judgment of each type, the methodist will retain the general principle judgment and revise his case judgment.

Kelly recommends that we think of these two methodologies as lying on a spectrum: near the far end of the spectrum is hyper methodism, according to which one should give almost no weight at all to one's judgments about cases; on the other end is hyper particularism, according to which one should give almost no weight at all to one's judgements about general principles; and in the middle is the view that judgments at the different levels of generality should be accorded equal weight.

Kelly's discussion of these views presupposes that one point on this spectrum is correct for all cases in which there's a conflict between one's case judgments and one's judgments about principles. It is this presupposition that I will take issue with. But first, let's get Kelly's argument on the table.

Kelly's first claim is that the correct principle of philosophical methodology is not very far over on the methodistic side of the spectrum. Kelly leaves open whether it is particularistic, or in the middle, or slightly on the methodistic side, but he claims that any kind of robust methodism is false, as a principle of philosophical methodology.

Kelly's defense of this claim is straightforward: often, we reject a general principle because of a single counterexample. That is, in cases of successful counter examples, we give up a judgment we have about a general principle because it conflicts with a single case judgment. This would not be possible if a principle of robust methodism were applied, since robust methodism assigns substantially more weight to general principles than to case judgments, and so a principle would not be given up just because it conflicts with a single case judgment. Thus, Kelly concludes that since we *should* sometimes give up a general principle because it conflicts with a single case judgment (as, for example, in the Gettier case), the correct principle of philosophical methodology is either close to the middle of the spectrum, or particularistic. Certainly, claims Kelly, any kind of robust methodism is false. And this result, says Kelly, is all that is needed to mount a Moorean response to the skeptic.

Kelly first notes that the skeptical argument relies crucially on a general principle about knowledge or evidence. Moreover, he says, anyone who accepts the skeptical argument will have to give up a very large number of judgments about cases, like the judgment that I know that I have hands,

that I know there is a table before me, etc. When we are confronted with the skeptic's argument, then, we are forced to choose between a general principle, on the one hand, and very large number of judgments about cases, on the other. According to any methodological principle that gives case judgments and principle judgments roughly equal weight (or more weight to case judgments), the recommended course of action is clear: give up the general principle in favor of the case judgments. If so, then the skeptical argument cannot overturn our common sense beliefs about the extent of our knowledge.<sup>18</sup>

The only way out of this conclusion, according to Kelly, would be to adopt a position of hyper methodism. (It is this claim that I will disagree with.) Kelly rightly notes that this option is highly undesirable, since a hyper methodist could never give up a general principle in the face of a single counter example.

As mentioned briefly above, I disagree with the presupposition implicit in Kelly's argument that one point on the spectrum of methodologies he describes must be appropriate for every case of conflict between general principles and case judgments. I agree with Kelly that in some cases, a general principle should be rejected because it conflicts with a single case judgment. But, it simply doesn't follow that one must never retain belief in a general principle that conflicts with many case judgments. The rational response to a conflict of this sort may depend on details of the case that go beyond the different numbers of types of judgments that one has. In some cases, it is rational to give up a general principle because it conflicts with one's judgments about cases. Kelly provides some examples of this sort. However, there are other cases in which rationality requires one to give up one's judgments about cases because they conflict with a general principle. I will provide some examples of this sort below.

Kelly's argument commits him to the view that in any case of conflict between many case judgments and a single general principle judgment, one must always reject the general principle judgment. Call this Kelly's Claim. I don't think Kelly's Claim is correct. I will now present two cases in which the rational response to a conflict of this kind is to give up the case judgments and retain the general principle judgement. If I am right about these cases,

<sup>18</sup> Nelson Goodman, in chapter three of *Fact, Fiction, and Forecast*, presents a view concerning inductive skepticism that, I think, is essentially the same as Kelly's argument against the external world skeptic. According to Goodman, we discover what a term means by finding a general principle about its meaning according to which the cases to which the term is *actually* applied just are the cases to which the term is *properly* applied. This is how we define "tree": we look for a general principle that captures what is in common between all the objects to which we actually apply the term "tree." Likewise, says Goodman, to define "valid induction," we should look for whatever it is that is in common between all of the arguments to which we actually apply the term "valid induction." Clearly, if we follow this procedure, we will never get the result that there are no valid inductions. So, if this procedure is correct, the inductive skeptic is wrong. The objections I will present against Kelly's Claim work equally well against this interpretation of Goodman's view. For example, if we were to apply Goodman's procedure in figuring out what "justified certainty" means, we would never reach what I will argue is the correct conclusion.

then Kelly's Claim is false, and so Kelly's argument for Mooreanism will not go through.

My first example is taken from the philosophy of probability. My second example is strongly parallel to the skepticism case, and thus is particularly instructive in this context.

First, consider someone who tends to commit the gambler's fallacy. If he sees a fair coin land heads many times in a row, he judges that the coin is more likely to land tails than heads on the next toss; if the coin lands tails many times in a row, he judges that heads is more likely than tails on the next toss; etc. Suppose that this person then takes a philosophy of probability class, and encounters the principle "Independent tosses of a fair coin are equally likely to come up heads." This principle strikes him as highly plausible. Then, however, it is pointed out to him that if he accepts this principle, he should no longer judge that heads is more likely than tails after a long series of tails (and likewise for the other gambler's fallacy judgments he is apt to make).

According to Kelly's Claim, the only rational response for this person is to regard the principle as false, since it conflicts with many of his case judgments. But this is clearly wrong: it may well be that the rational response for him is to give up the case judgments. After all, if this person reflects on the principle, he may come to see why it's true, and come to realize that his previous case judgments were mistaken and confused. In such a case we would want to say that the rational thing for this person to do is to retain belief in the general principle, and let this principle guide his case judgments. But this is exactly what Kelly's Claim says he must not do.

I think there are many examples of this sort: cases in which many of one's initial case judgments were based on conceptual confusion, which can be cleared up when one reflects on a conflicting (but true) general principle. Many of the other examples from the heuristics and biases literature could be used to make the same point.

I'll now describe my second counterexample to Kelly's Claim. Consider someone who has never taken a philosophy class, and who has never taken the time to reflect on the epistemic status of ordinary, day-to-day beliefs. Let's call this person Ordinary. Now imagine that a philosopher goes and has a chat with Ordinary. They are sitting on a sofa in front of a table, on which there lies a book, in plain view, in natural light, etc. Upon chatting with Ordinary, the philosopher discovers that he believes many things, including the following: he believes that there's a book on the table and he believes that the sun will rise tomorrow. Moreover, the philosopher discovers that Ordinary is certain of these things, and—crucially—that he takes himself to be justified in being certain of them. Having never heard of the fanciful skeptical hypotheses familiar to philosophers, Ordinary takes himself to have definitively ruled out all possibilities in which these propositions are false, and so takes himself to be fully justified and rational in being certain that they are true.

Suppose, however, that the philosopher goes on to describe to Ordinary certain skeptical scenarios, involving evil demons who trick people into

believing that there are books in front of them when there really aren't, counter inductive worlds in which one remarkable day the sun fails to rise, etc. The philosopher goes on to explain to Ordinary that, since such scenarios are possible, many of Ordinary's judgments of the form "I am justified in being certain that P" conflict with the principle "If there is a possible scenario in which not-P that one cannot definitively rule out, then one is not justified in being certain that P." Now that he is aware of this contradiction, Ordinary must choose between his belief in this highly plausible general principle and many of his beliefs of the form "I am justified in being certain that P."

How should Ordinary revise his beliefs in this case? I think most epistemologists would agree that he should retain belief in the general principle and give up the case judgments. Of course, most would also say that he should still think that he knows these propositions, or at least that he's justified in believing them. But he should no longer think that he's justified in being certain of them, since he no longer takes himself to have ruled out all possible scenarios in which they are false.

This is another case in which a large number of one's initial case judgments were mistaken as a result of conceptual confusion which can be cleared up by reflection. In this case, we might say that Ordinary was led into these mistakes by a limited imagination: he failed to realize that scenarios like the skeptical scenarios were possible; they had simply never occurred to him. This failure to recognize the possibility of scenarios of a certain sort led him to misapply his own concept of justified certainty. So this is a case in which Ordinary should, upon realizing this mistake, retain belief in the general principle about certainty and let its proper application guide his case judgments.

These two counterexamples to Kelly's Claim illustrate a general way in which it goes wrong. Kelly's Claim does not allow for the possibility that one systematically misapplies one's own concepts due to conceptual confusion.<sup>19</sup> But surely this is possible, as these two cases illustrate. It is possible to systematically misapply one's concept of comparative probability, and it is possible to systematically misapply one's concept of justified certainty. The skeptic claims that, just as it is possible to systematically misapply these concepts, it is possible to systematically misapply our concept of knowledge. And this, contends the skeptic, is exactly what we have in fact done.

In comparing the skeptical case to the two cases just discussed, I think it is especially helpful to focus on the certainty case. This is because the response to the conflict in the certainty case that is widely agreed by epistemologists to be the rational one is highly structurally analogous to what the skeptic takes to be the rational response in the case of knowledge. In both cases, there is some general principle about the conditions under which one stands in a certain kind of epistemic relation to a proposition. In the one case, it is a principle

<sup>19</sup> Strictly speaking, Kelly's Claim does not allow for the possibility that one rationally believes that one has systematically failed to misapply one's own concept.

about when one is justified in being certain of something; in the other case, it is whatever principle about knowledge is employed in the skeptical argument. In both cases, one initially judges, of many of the propositions that one believes, that one *does* stand in that epistemic relation to them. In fact, it may well be almost the same set of propositions in each case. Most epistemologists would recommend revising one's belief, about each such proposition, that one is justified in being certain of it. The skeptic recommends revising one's belief, about each such proposition, that one knows it. The two cases of belief revision are, structurally, almost exactly the same.

We have just seen that the certainty case blocks Kelly's anti-skeptical argument by showing that the methodological principle he relies on—Kelly's Claim—is not true. In general, it is a consequence of the close structural similarity of the certainty case and the knowledge case that it will be extremely hard for *any* general principle of philosophical methodology to yield the right result about the certainty case—namely, that one should give up one's initial case judgments—while also yielding the result that one should retain one's initial case judgments about what one knows.

It's time to take stock. Kelly endorses the methodological claim that one should always give up a general principle when it conflicts with many case judgments. If this is true, then, when confronted with a conflict between a general philosophical principle and common sense judgments about particular cases, one should give up the philosophical principle and retain belief in the common sense judgments. However, I have argued that Kelly's methodological principle fails because it does not allow for the possibility that one has systematically misapplied one's concepts. I used two cases—the gambler's fallacy case and the certainty case—to illustrate that systematic misapplication of concepts is indeed possible.

## 7. CONSERVATISM AND REFLECTIVE EQUILIBRIUM

In this section I will discuss a version of reflective equilibrium that emphasizes *minimal change* as an important aspect of rational belief revision—not just in philosophy, but belief revision of any kind. I will focus on the version elaborated and defended by Gilbert Harman, but note that Lewis 1999, Pryor 2000, and others also take conservatism to be an important part of proper philosophical methodology.

In the section entitled “Philosophical Skepticism” of his paper “Skepticism and Foundations,” Harman describes the basics of his version of reflective equilibrium and argues that it has the result that we should retain belief in the common sense proposition that we know we have hands and give up one of the skeptic's premises.

Harman begins by considering a person who has the usual beliefs about the external world, and about the extent of his knowledge of the external world, and who also believes (implicitly, perhaps) the premises of the skeptical

argument. Suppose this person then discovers the contradiction between his beliefs about the extent of his knowledge and the skeptical premises that he accepts. How should this person revise his beliefs in order to avoid this inconsistency?

Harman's principle says that "the way to resolve such conflicts is by finding the minimal adjustment in *S*'s beliefs and methods that provides the most simple and coherent upshot."<sup>20</sup> Upon discovering a contradiction in one's beliefs, one must choose to update to just one of the many possible complete systems of belief that resolve this contradiction in some way. According to Harman's principle, you should adopt whichever one of these complete systems of belief best balances minimal change, coherence, and simplicity.

Now consider the particular case at hand, in which one has discovered a conflict between the skeptic's premises and one's beliefs about the extent of one's knowledge. Retaining the skeptical premises requires giving up a vast number of beliefs about the external world, and about one's knowledge of it. This would involve a very large change in one's system of belief. Giving up one of the skeptical premises, however, would require only a very small change in comparison; all of one's beliefs about the external world, and one's knowledge of it, would remain intact. Thus, Harman's principle of belief revision, with its emphasis on minimal change, yields the result that one ought to give up one of the skeptical premises.<sup>21</sup>

Lewis expresses the same general thought in the following quote from "Elusive Knowledge:" "We are caught between the rock of fallibilism and the whirlpool of skepticism. Both are mad! Yet fallibilism is the less intrusive madness. It demands less frequent corrections of what we want to say. So, if forced to choose, I choose fallibilism."<sup>22</sup> I'll focus on Harman's version of the idea, though, since it's developed in more detail.

I will argue that Harman's principle of rational belief revision is not a good one. First, note that the counterexamples to Kelly's Claim, discussed in the previous section, also constitute counterexamples to Harman's principle. In these counterexamples, it is rational to revise a great many of one's initial case judgments because they conflict with a single general principle. Because there are so many case judgments that must be revised, rejecting the general principle instead would have constituted a more minimal change. So, Harman's principle shares the flaw identified earlier in Kelly's principle: it does not allow for the possibility of systematic misapplication of one's concepts.

However, I want to develop and bring out another, perhaps deeper, problem with Harman's principle of rational belief revision.

<sup>20</sup> Harman (2003, 10).

<sup>21</sup> In fact, Harman needs to make a few further assumptions in order to reach this conclusion. It must be assumed that giving up the skeptical principle does not result in large losses in simplicity and coherence; and, it must be assumed that the option of retaining the skeptical principle and the claims to knowledge, and giving up the belief that they conflict, would not be acceptable.

<sup>22</sup> Lewis (1999, 419).

As noted above, while Harman takes his principle to correctly describe rational revision of beliefs in philosophy, he does not think it is limited to beliefs in this area. Indeed, Harman takes his principle to apply to all rational belief revision. I will present a counterexample to Harman's view involving a case of evidence by testimony to bring out the second general problem with Harman's principle. Ultimately, I will argue that Harman's principle fails to be appropriately sensitive to the relations of epistemic dependence that obtain between some beliefs. I will claim that the nature and number of beliefs that depend epistemically on a belief *Q* are irrelevant to whether *Q* should be given up. However, whether Harman's principle recommends giving up belief in *Q* is sometimes determined in part by the nature and number of beliefs that depend on *Q*. This, I think, is the central problem with Harman's principle.

I'll start by describing a case that does not itself constitute a counter example to Harman's principle, but which will help me set up the counter example.

*Case 1:* Imagine that you have lived your entire life in a remote, isolated village, and know nothing about the world beyond the village borders. One day you encounter a visiting stranger, Alice, who tells you she's spent her life traveling around the world, including a place called "Costa Rica," and proceeds to tell you about the beautiful birds she saw there. Naturally, you believe what she tells you. Later, however, you meet up with one of your best buddies—Bert. Bert has an uncanny knack for being able to tell when people are lying. Time and again he's told you that someone was lying about something, and independent investigation proved him right. This time, Bert tells you that Alice was lying—in fact, she has never been to Costa Rica.

It is clear that in this situation you should believe Bert. After all, Bert has an excellent track record of detecting lying, and you only just met Alice. Now consider a modified version of this case. Case 2 (below) is exactly like case 1, with the following exception:

*Case 2:* Alice didn't just tell you a few things about Costa Rica. Rather, she has told you many stories about her travels all around the world. You love listening to these stories, and you have spent most of your evenings listening to them. As a result, you have accumulated a vast and detailed system of beliefs about the world beyond the village border, all on the basis of Alice's testimony. It is only after this has been going on quite a while that Bert tells you that Alice has been lying the whole time.

I think it's equally clear that you should believe Bert in case 2 as well. After all, as before, Bert has an incredible track record; and, the mere fact that Alice has told you many things does not make any particular one of them more credible.

However, Harman's principle says otherwise. In case 2 (but not case 1) believing Bert would require you to give up a vast number of other beliefs that you have—you would have to give up all your beliefs about the nature of



the world outside your village. This constitutes quite a substantial change in your overall system of belief. Such a substantial change would not be required if you were to believe instead that Bert just happened to be wrong in this particular case. This would be quite a minimal change, overall, since you could retain all the beliefs you formed on the basis of Alice's testimony. So, Harman's principle says you should retain belief in everything that Alice told you. But clearly this is not the rational response to hearing what Bert said. You should believe Bert, even though this requires you to give up all of your beliefs about what lies beyond the village border.

Before explaining what general problem I think this illustrates with Harman's view, I will consider and respond to a possible response from a defender of Harman's principle of belief revision.

Harman might respond that in fact, his principle does not have the consequence that I claim it has. I claimed that Harman's principle has the consequence that, in case 2, you should retain believe in what Alice told you, since this makes for a more minimal change. But Harman's principle says that minimal change is just one criterion, to be balanced with the other criteria of simplicity and coherence. While there doesn't seem to be any difference in simplicity between the two belief systems, there may seem to be a difference in coherence. It will be helpful to describe in more detail the two complete belief systems you're choosing between. I've named them BERT and ALICE:

BERT: Bert is right that Alice was lying to me; after all, he's always been right in the past. So, I cannot trust anything that Alice told me, and must give up all the beliefs I had formed on the basis of her testimony.

ALICE: Alice has been telling me the truth the entire time. Although Bert is generally right, he was wrong in this case.

Harman might claim that ALICE, while a more minimal change than BERT, is less coherent than BERT. After all, on belief system ALICE, you believe that Bert is generally reliable, but you also believe that he was wrong in this case, even though you don't also believe that there is some particular feature of this case in virtue of which he was likely to make a mistake. Surely this collection of beliefs doesn't hang together very well, and could be seen as making for a slight incoherence in belief system ALICE.

I think this putative problem can be circumvented. A slight modification of ALICE will get rid of the minor incoherence. Consider ALICE\*:

ALICE\*: Alice has been telling me the truth the entire time. Although in general when Bert tells me someone is lying, that person really is lying, I can see that Bert is jealous of all the time I'm spending with Alice, and he told me she is lying only because he hopes it will have the effect of my spending less time with her. I couldn't explain to someone in words exactly what it is that makes me believe he's jealous, but I can just tell he is by the way they interact.

ALICE\* does not have any of the mild incoherence that might have been found in ALICE; and, it is still a much more minimal change than BERT. So, even if Harman's principle does not yield the result that one should adopt belief system ALICE, it does say that rather than adopting belief system BERT, you should revise your system of beliefs to ALICE\*, since that would constitute a much more minimal change, and it is no less simple or coherent.

So, for example, consider someone who, prior to hearing Bert's claim that Alice is lying, did not suspect that Bert was jealous, and who has no good evidence that Bert is jealous. However, this person is loath to make any very serious changes to his belief system, and whenever he encounters evidence against many of his beliefs, he has a tendency to come up with some way of explaining the evidence away, rather than coming to terms with the real implications of the evidence (of course he would not describe himself that way). So, when this person hears Bert's claim that Alice has been lying, he immediately begins to suspect that there's something fishy going on, and eventually convinces himself that Bert is jealous of Alice, etc.—in short, he comes to accept belief system ALICE\*. According to Harman's principle, this person is revising his beliefs exactly as he should be. But this is the wrong result. The rational response to Bert's testimony is not to continue to believe Alice, but rather to give up belief in everything Alice said.

Now that I've argued that this case does indeed constitute a problem for Harman's principle, I'll move on to the diagnosis: what deeper problem with Harman's principle is illustrated by this counterexample?

The core of my diagnosis will be as follows. There is a relation of epistemic dependence that holds between some beliefs (I will say more about epistemic dependence in a moment). I claim that beliefs that depend epistemically on some other belief Q are irrelevant to whether or not Q should be given up. That is, how many beliefs there are that depend on Q, and what the contents of those beliefs are, is irrelevant to whether belief in Q should be maintained. But whether Harman's principle recommends giving up belief in Q is sometimes determined in part by beliefs that depend on Q. This is the problem with Harman's principle that is brought out by the counterexample I gave.

In order to further explain and defend this diagnosis, I need to give some explanation of the notion of epistemic dependence. There are several formal properties of the dependence relation that can be noted at the outset. First, dependence is relative to an agent; it may be that P depends on Q for one agent but not for another. Second, the dependence relation is never symmetric. If P depends on Q, then Q does not depend on P (relative to the same agent, of course).<sup>23</sup>

<sup>23</sup> Some epistemologists (perhaps some coherentists) may disagree with me about this: they may think that some cases of dependence *are* symmetric. That's OK for me, as long as the coherentist is willing to grant that *some* cases of dependence are asymmetric, and, in particular, that my belief that there are beautiful birds in Costa Rica asymmetrically depends on my belief that Alice has been telling the truth. Instead of claiming that beliefs that depend on Q are never

Now for a more substantive characterization of dependence. The intuitive idea is that P depends on Q for an agent S when S's belief that P is based (at least in part) on S's belief that Q, and when Q (perhaps in conjunction with other beliefs that S has) does indeed justify P; P derives its justification (at least in part) from Q.

Relations of epistemic dependence are commonplace. My belief that the chicken I just ate had not gone bad depends on my belief that I took the chicken meat out of the freezer and into the fridge only a day or two ago (and not, say, several weeks ago). My belief that the temperature is 84 degrees depends on my belief that my thermometer reads "84." My belief that the butler did it depends on my belief that the butler's fingerprints were found at the scene of the crime.

Moreover, your belief that there are beautiful birds in Costa Rica depends on your belief that Alice is telling the truth. Harman's principle says that you should retain your belief that Alice is telling the truth because giving it up would require you to give up your belief that birds in Costa Rica are beautiful, and all other such beliefs. On Harman's principle, part of what makes it the case that you are justified in holding on to your belief that Alice is telling the truth is that you have many other beliefs that depend on it, and which you would have to give up with it. But that is to make beliefs that depend on Q relevant to whether or not belief in Q should be retained. In particular, it is to make beliefs that depend on your belief that Alice is telling the truth relevant to whether you should continue to believe that Alice is telling the truth. And it is my contention that beliefs that depend on Q are never relevant to whether belief in Q should be retained. Whether you have any beliefs that depend on Q, and if so, what beliefs they are, is irrelevant to whether or not it would be epistemically rational for you to retain belief in Q. The problem with Harman's principle is that it is not consistent with this fact.

Let's take stock. I have argued that there are two main problems for Harman's principle of belief revision. The first is that, like Kelly's Claim, it does not allow for the possibility of systematic misapplication of concepts. The second is that it allows the number and nature of beliefs that depend on Q to be relevant to whether belief in Q should be retained. Thus, Harman's principle is false, and so, like Moore's plausibility principle and Kelly's Claim, it cannot be used to motivate the idea that philosophy can't overturn common sense.

## 8. SUMMARY AND CONCLUDING REMARKS

It has become popular to think of common sense as an oracle to which the philosopher must always defer. If a philosopher's theory turns out to conflict with common sense, the philosopher is taken to have overstepped her bounds

relevant to whether Q should be given up, I would then re-state my claim as follows: beliefs that depend *asymmetrically* on Q are never relevant to whether Q should be given up. The argument against Harman goes through just as well this way.

and is expected to retreat. It has been my aim in this paper to convince the reader that this conception of philosophy is untenable. Philosophical argument is perfectly capable of undermining our ordinary, pre-theoretic view of the world.

In the first half of the paper, I provided a positive argument for my central claim that philosophy can overturn common sense. In particular, I argued that if (as my opponents agree) science can overturn common sense, then so can philosophy. In the second half of the paper, I objected to the main motivations philosophers have had for taking the opposing view, namely that philosophy is not powerful enough to overturn common sense. These motivations turned out to rely on faulty theories of philosophical methodology.

One consequence is that we cannot simply dismiss out of hand philosophical arguments—like arguments for skepticism—that target common sense claims. We cannot know in advance that our ordinary beliefs will stand fast in the face of such arguments; only careful and detailed consideration of these arguments can reveal whether or not they succeed. This brings a heightened sense of importance and urgency to philosophical inquiry. Nothing less than our most basic and central beliefs are at stake.

#### REFERENCES

- Chisholm, Roderick (1973) *The Problem of the Criterion* Milwaukee: Marquette University Press.
- Christensen, David (2007) "The Epistemology of Disagreement: The Good News" *Philosophical Review* 116: 187–217.
- Conee, Earl (2001) "Comments on Bill Lycan's Moore Against the New Skeptics" *Philosophical Studies* 103 (1): 55–59.
- Dorr, Cian (2002) "The Simplicity of Everything," Princeton University doctoral dissertation.
- Elga, Adam (2007) "Reflection and Disagreement" *Nous* 41 (3): 478–502.
- Fine, Kit (2001) "The Question of Realism" *The Philosophers' Imprint* 1 (2): 1–30.
- Goodman, Nelson (1983) *Fact, Fiction, and Forecast* Cambridge, Mass: Harvard University Press.
- Gupta, Anil (2006) *Empiricism and Experience* Oxford: Oxford University Press
- Harman, Gilbert (2003) "Skepticism and Foundations," in Steven Luper (ed.) *The Sceptics: Contemporary Essays* Aldershot: Ashgate Publishing: 1–11.
- Hegel, G. W. F. (1977) *Phenomenology of Spirit* Oxford: Oxford University Press.
- Kahneman, D., Tversky, A., and Slovic, P. (1982) *Judgment Under Uncertainty: Heuristics and Biases* New York: Cambridge University Press.
- Kant, Immanuel (2008) *Prolegomena to Any Future Metaphysics* New York: Cambridge University Press.
- Kelly, Thomas (2005) "Moorean Facts and Belief Revision or Can the Skeptic Win?" in John Hawthorne (ed.) *Philosophical Perspectives: Epistemology* 19: 179–209.

- (2008) "Common Sense as Evidence: Against Revisionary Ontology and Skepticism," in Peter French and Howard Wettstein (eds.) *Midwest Studies in Philosophy: Truth and Its Deformities* 32: 53–78.
- Lewis, David (1973) *Counterfactuals*, Cambridge, Mass: Harvard University Press.
- (1999) "Elusive Knowledge," in David Lewis (ed.) *Papers in Metaphysics and Epistemology* Cambridge: Cambridge University Press, 248–261.
- Lycan, William (2001) "Moore Against the New Sceptics" *Philosophical Studies* 103 (1): 35–53.
- Moore, G. E. (1962) "A Defense of Common Sense," in G. E. Moore (ed.) *Philosophical Papers* London: Collier Books.
- (1962) "Four Forms of Skepticism" in G. E. Moore (ed.) *Philosophical Papers* London: Collier Books.
- (1962) "Proof of an External World" in G. E. Moore (ed.) *Philosophical Papers* London: Collier Books.
- Pryor, James (2000) "The Skeptic and the Dogmatist" *Nous* 34: 517–549.
- Rinard, Susanna (ms) "Reasoning One's Way out of Skepticism".
- Unger, Peter (1975) *Ignorance: A Case for Skepticism* Oxford: Oxford University Press.
- Van Inwagen, Peter (1990) *Material Beings* Ithaca: Cornell University Press.