Michael Huemer, "Direct Realism and the Brain-in-a-Vat Argument"

One of the advantages traditionally claimed for direct realist theories of perception over indirect realist theories is that the direct realist is able to avoid skeptical problems to which the indirect realist falls prey.[1] If the only things we are ever directly aware of are the ideas in our own minds, it is asked, then what reason do we have for thinking anything other than ideas exists? How do premises about ideas confirm propositions about physical objects? This is one sort of skeptical worry that the direct realist has an obvious *prima facie* advantage in dealing with.

However, there are other sorts of skeptical problems that direct realism does not seem to particularly help us with, and it is on one of those that I want to focus. Specifically, does the direct realist have an answer to brain-in-a-vat skepticism that is not available to the indirect realist? I claim that the answer is yes.

Before we are in a position to see that, we'll first have to review the brain-in-a-vat argument and explain the distinction between direct and indirect realism. After that, we'll need to look at two contemporary responses to the brain-in-a-vat argument to see why they fail as long as the direct realist account of perception is neglected. Then we'll be able to see how direct realism figures in the refutation of the skeptical argument.

1. Direct and indirect realism

Direct realism is often understood as the view that, in cases of normal perception, we are directly aware of something in the external world. This "something" could include external objects, events, or states of affairs; surfaces of external objects; and/or properties of external objects. Indirect realism is then characterized as the view that, in normal perception, we are only directly aware of internal (mental) phenomena, and we are *indirectly* aware of external phenomena, by means of our awareness of the mental phenomena. These internal, mental phenomena could include mental objects, states, events, and/or properties. So there can be different versions of indirect realism, according to what the theorist says about the nature of the mental phenomena that perception makes us aware of: the indirect realist might hold that what we are directly aware of are sense data, or states

Michael Huemer, "Direct Realism and the Brain-in-a-Vat Argument," *Philosophy and Phenomenological Research* 61 (2000): 397–413.

of "being appeared to" in certain ways, or some other sort of mental phenomena.

That is one traditional way of formulating the issue, but I'm going to use a slightly different formulation here. For our present purposes, what we need is an explicitly epistemological characterization of direct and indirect realism, a characterization in terms of the justification for perceptual beliefs. A person's belief that P is a *perceptual belief* if the (causal) explanation for why he believes that P is that he perceives that P. For instance, the explanation for why I believe that there is a pen here might be that I *see* that there is a pen here, which is a particular way of perceiving that there is a pen here, so I have the *perceptual belief* that there is a pen here. Now we can define indirect realism as the view that, at least in normal cases, perceptual beliefs about the external world are justified, but they depend for their justification on our being aware of certain mental phenomena. So for example, my belief that there is a pen here might depend for its justification on my awareness of a sense datum of a pen, or on my awareness of a state of being appeared to penishly, etc. *Direct realism* is the view that, at least in normal cases, our perceptual beliefs about the external world have justification that does not depend on our being aware of mental phenomena, or anything else that's not in the external world.[2]

2. The brain-in-a-vat argument

The brain-in-a-vat argument goes like this:

(1) If S is justified in believing P and P entails Q, then S is justified in believing Q.[3]
(2) I'm not justified in believing that I'm not a brain in a vat.
(3) That I have a body entails that I'm not a brain in a vat.
(4) Therefore, I'm not justified in believing that I have a body.

It is important to understand that the conclusion of this argument is not merely that I cannot be *absolutely certain* that I have a body. Such a conclusion would be accepted by many philosophers without much excitement. But the argument we're considering is much more interesting. It purports to establish that I don't even have good reason for believing that I have a body. And of course, the assumption is that if you can't justifiably believe that you have a body, then you can't justifiably believe much of anything about the physical world.

A note about the phrase "justified in believing": "S is justified in believing P," as I use the phrase, does not entail that S actually believes P. Rather, it just means that S has available an adequate source of justification for P. S might fail to actually accept some of the propositions that are justified for him, either because S is epistemically timid or because he hasn't noticed the justification he has available. To illustrate the distinction, let's suppose that S knows that P and he also knows that $(P \supset Q)$; however, S has not yet put together these two beliefs and noticed their logical consequence, so S doesn't actually believe that Q. (Presumably, this sort of thing takes place frequently, because we are often aware of certain axioms without being aware of most of the theorems derivable from them.) Nevertheless, S *has available* an adequate justification for Q, insofar as P and $(P \supset Q)$ would provide adequate reason for believing Q, so I would say that S

is "justified in believing Q."[4] In general, S is justified in believing P whenever S is in possession of some evidence that, if properly deployed, would lead S to believe P and be justified in doing so.

3. Why direct realism may be irrelevant to the argument

At first glance, it is not obvious how the issue between direct and indirect realism is relevant to this skeptical argument. One could say that, if direct realism is true, it follows that the skeptical argument is unsound, since direct realism as we have defined it involves the claim that we are justified in believing propositions about the external world. But it is equally true that, if *indirect* realism is true, then the skeptical argument is unsound, since indirect realism also involves the claim that we are justified in believing propositions about the external world. So far, no advantage for direct realism is evident. And neither of these observations constitutes a response to the skeptical argument; the skeptic will simply reject both forms of realism as I have formulated them.

One might try arguing that the brain-in-a-vat scenario, as usually described, is impossible, in the sense that the brain in a vat could not really have the same kind of experiences we have. We have experiences of *perceiving* (and hence, being directly acquainted with) external objects. But the brain in the vat doesn't have any *perceptions*. The brain in the vat only has *hallucinations*. According to one faction of direct realists, perceptions and hallucinations do not really share anything in common; they are merely states that *seem* alike to the subject at the time, but this does not show that they actually are intrinsically alike.[5] Thus, John Hyman writes:

> [T]he causal theory is still committed to the Cartesian illusion that "the ordinary notion of perceiving" is a composite notion, which can be divided into its purely mental and its physical components, each of which can exist without the other . . . Against this view I have argued, first, that perceptual experience and illusion are not two species of the same psychological genus and hence the concept of perceiving cannot be dismembered in this way; and second, that there is no epistemological reason for trying to dismember it, since the foundations of empirical knowledge are propositions stating what the speaker perceives.[6]

Supposing that this view is right—namely, that there is no mental state common to both hallucination and veridical perception—does it furnish us with an answer to the brain-in-a-vat argument?

It certainly provides us with a way to object to one way the brain-in-a-vat scenario is commonly described—namely, as a situation in which a being would have the same experiences we presently have, but most of his beliefs about the external world would be false. However, the important question is really whether hallucinations and veridical perceptions can be subjectively indistinguishable (that is, indistinguishable for the subject, at the time). If the direct realist holds that there can not be hallucinations that are subjectively indistinguishable from perceptions, then his theory is just empirically false. We have every scientific reason to believe that if a brain were electrically stimulated in the way described

in the brain-in-a-vat scenario, its hallucinations would be subjectively indistinguishable from the perceptions of a normal person. But if the direct realist admits that hallucinations can be subjectively indistinguishable from perceptions, then the skeptic can merely rephrase his challenge: how do we know that what we are actually doing is perceiving things, rather than merely hallucinating? Whether perceptions and hallucinations have a common component isn't what matters; what matters to the skeptic's question is whether we have a way of distinguishing perceptions from hallucinations in our own case. As long as we can't tell the difference, a skeptical worry will remain. All the direct realist has accomplished, so far, is to get the skeptic to reformulate his question.

4. Two contemporary responses

Premise (1) of the skeptical argument is called "the Closure Principle"—the principle that the set of propositions one is justified in believing is closed under entailment. This principle is highly plausible intuitively, but some epistemologists have challenged it.[7] Fred Dretske cites the following case. Imagine you're at the zoo. In a pen clearly marked "zebras," you see some black and white striped equine animals. It seems that you have good reason to believe that those animals are zebras. Surely their zebra-like appearance counts strongly in favor of their being zebras, as does their being in the zebra pen at the zoo. Now, their being *zebras* entails that the animals are not *mules* that have been cleverly disguised by the zoo authorities to look like zebras. Are you justified in believing that the animals are not cleverly disguised mules? Dretske says no:

> If you are tempted to say "Yes" to this question, think a moment about what reasons you have, what evidence you can produce in favor of this claim. The evidence you *had* for thinking them zebras has been effectively neutralized, since it does not count toward their *not* being mules cleverly disguised to look like zebras.[8]

Dretske views this as a counter-example to the Closure Principle: you are justified in believing that the animals are zebras, that they are zebras entails that they are not cleverly disguised mules, but you're not justified in believing that they are not cleverly disguised mules.

What can we say against this? Can we define the intuition behind the Closure Principle? The Closure Principle holds that when S is justified in believing P and P entails Q, S is justified in believing Q. There are at least two reasons why this might be the case. One reason would be that *the very same justification* S has for P also counts as justification for Q—i.e., whatever evidence supports P would also support Q when Q is a logical consequence of P. If that were generally true, then it would follow that the Closure Principle is true. Dretske's example effectively refutes *that* principle. Dretske's example shows that what is evidence for P need not be evidence for every logical consequence of P. The fact that the animals in the pen look like zebras is evidence that they are zebras, but it is not evidence that they aren't mules cleverly disguised to *look like zebras*.

However, another reason for believing the Closure Principle is this: if S is justified in believing P and P entails Q, then *P*, itself, constitutes an adequate

reason for believing Q. The idea is simply that deduction is an epistemically permissible way to expand your corpus of beliefs. This idea is probably the real source of the intuition in favor of closure. What Dretske says about his zebra-in-the-zoo case does not address this idea; what he says is only that the evidence you have for thinking the animals are zebras is not evidence against their being cleverly disguised mules. That much seems clear. But Dretske doesn't explain why the fact that *the animals in the pen are zebras* wouldn't be a sufficient reason for thinking that they're not cleverly disguised mules, given that you justifiably believe that the animals are zebras.

Peter Klein has pressed the above point.[9] However, Klein argues that, even though the above reply enables the skeptic to defend premise (1) against Dretske's attack, it nevertheless leaves the skeptic with a problem in defending premise (2). To defend premise (2), what the skeptic has to argue is that I have no available source of justification for the proposition that I'm not a brain in a vat. In defending the Closure Principle, we have just said that when P entails Q and P is justified, P is itself an adequate source of justification for Q—deductive arguments are a source of justification. So the skeptic must argue that, among other things, I don't have *that* kind of justification for believing I'm not a brain in a vat. Since, as premise (3) now assures us, the proposition that I have a body would provide just that sort of adequate justification for thinking I'm not a brain in a vat, the skeptic would have to argue that I don't have that proposition in particular available as a source of justification for thinking I'm not a brain in a vat.

What this means is that, in order to establish premise (2), the skeptic would first have to establish that I'm not justified in believing that I have a body, since that belief, if justified, would be one adequate source of justification for the claim that I'm not a brain in a vat. But that I'm not justified in believing that I have a body is just the conclusion of the argument. So it seems that the skeptical argument begs the question—one of its premises can't be established unless the conclusion is established first.[10]

To put the point another way:[11] suppose I start out thinking that I'm justified in believing I have a body, and the skeptic then proposes to argue me out of this position. He starts by informing me that the Closure Principle is true, because it is epistemically permissible to add to your body of beliefs the deductive consequences of any of your justified beliefs. The skeptic then asserts that I have no available justification for believing I'm not a brain in a vat. I naturally reply, "Yes I do, because I justifiably believe I have a body, which entails that I'm not a brain in a vat, and you just told me that it is epistemically premissible to add to my belief system the deductive consequences of any of my justified beliefs." What will the skeptic say? Why is this not an adequate source of justification for thinking I'm not a brain in a vat? Because I'm *not* justified in thinking I have a body? But that's just the conclusion the skeptic is trying to establish; I'm not going to grant that off hand. So the skeptic needs some other argument for the claim that I'm not justified in thinking I have a body. But if he has such an argument, then he doesn't need to use the brain-in-a-vat argument to begin with, because he would have an independent argument for the same conclusion.

5. What's wrong with these responses?

In short, Dretske's response to the argument is this: Okay, I don't know whether I'm a brain in a vat, but that doesn't matter; I still know that I have a body. And Klein's suggestion is this: Suppose we grant the Closure Principle. Then the skeptic's claim that I'm not justified in believing I'm not a brain in a vat just begs the question.

In spite of what we have said above in the way of philosophical analysis, I think intuition still balks at these responses. It seems as if there must be something wrong with them. It doesn't seem right that I can just admit that I don't know whether I'm a brain in a vat and continue to go on claiming to know all the things I have hitherto thought I knew. But nor does it seem right that the fact that I have two hands could be an adequate proof that I'm not a brain in a vat.

Let's try to articulate why the responses seem wrong. Consider the following two, possibly analogous cases:

> *Case (i) (the courtroom case)*: Imagine that S is on trial for murder. The prosecution offers as evidence the fact that S's blood was found at the scene of the crime along with the victims' blood. The best explanation for this, they say, is that S cut himself while stabbing his victims. The jurors find this argument initially persuasive. However, the defense attorney offers an alternative hypothesis: perhaps S is innocent, and the blood was planted at the crime scene by overzealous police officers seeking to frame S.

We can imagine how jury members might react to the defense hypothesis. Some jurors might feel that, being unable to rule out the alternative hypothesis, they should acquit S. Jurors arguing for a conviction might argue that the defense hypothesis should be rejected because it requires an improbable conspiracy on the part of the police department, because the police had no motive to frame S, and so on. But one thing that a jury member could not be expected to say is the following: "Okay, I agree that we have no reason for rejecting the defense hypothesis. But I still think we should convict S anyway, because we know he did it." Another thing that a juror would probably not come up with is this: "The defense attorney's claim that we can't rule out his hypothesis begs the question, because if we know S is guilty, then we *can* rule out the defense hypothesis."

The first of these unsatisfactory responses parallels Dretske's response to the skeptic. The second parallels Klein's response. If either of these responses were offered, they would probably be met with looks of puzzlement from the other jury members.

> *Case (ii) (the scientific case)*: Two scientists are arguing over the interpretation of quantum mechanics. Physicist A proposes the Copenhagen interpretation, noting that it accounts for a number of weird experimental results. The Copenhagen interpretation is the received view. Physicist B then proposes Bohm's interpretation of quantum mechanics, which is incompatible with the Copenhagen interpretation, noting that Bohm's theory accounts for all of the same experimental results. Now A might be expected to object that Bohm's theory conflicts with relativity, or that it is

somehow less parsimonious than the Copenhagen interpretation. But one thing A would probably not say is the following: "Okay, I agree that I can't rule out Bohm's theory; for all I know, that may be the right interpretation. But nevertheless, I still know that the Copenhagen interpretation is correct." Nor could we expect A to resort to the following objection: "Your claim that I can't rule out Bohm's theory begs the question, because if I know the Copenhagen interpretation is right, and Bohm's theory conflicts with the Copenhagen interpretation, then I *can* rule out Bohm's theory."

Again, both of these would strike us as illogical replies; yet they are, respectively, analogous to Dretske's and Klein's responses to the brain-in-a-vat argument. If Dretske's or Klein's response to the brain-in-a-vat argument is correct, then one of these absurd replies should be correct in the courtroom case and the scientific case.

Dretske, of course, would challenge the analogy. It is not his view that, in order to know something, one *never* needs to rule out alternative possibilities. Rather, his view is that there are certain kinds of alternatives that one needs to rule out (call them the "relevant alternatives") in order to know something, and there are other kinds of alternatives that one does not need to rule out (the "irrelevant alternatives").[12] Dretske would claim that the brain-in-a-vat hypothesis is an irrelevant alternative, but the defense hypothesis in the courtroom case and Bohm's theory in the scientific case are each *relevant* alternatives. Of course, this claim remains only a promissory note until it is explained what makes an alternative "relevant." According to Dretske, an alternative is relevant only if it is *genuinely possible*, in a certain sense:

> [T]he difference between a relevant and an irrelevant alternative resides, not in what we happen to *regard* as a real possibility (whether reasonably or not), but in the kind of possibilities that actually exist in the objective situation.[13]

Dretske doesn't give a precise analysis of the sense of "possible" he is invoking here, but his discussion makes it clear that it is a sense stronger than logical possibility, stronger than physical possibility, and non-epistemic.[14] Whether something is genuinely possible is supposed to be independent of our beliefs, evidence, and/or knowledge.

Dretske might argue that the brain-in-a-vat hypothesis is an irrelevant alternative, on the grounds that it is not, in his sense, genuinely possible for me to be a brain in a vat (perhaps because no one possesses the technology for keeping a disembodied brain alive, nor for stimulating it in the right ways). As a result, it is not a condition on our knowing about the external world that we rule out the brain-in-a-vat hypothesis. However, Dretske would have difficulty distinguishing this case from the scientific case. In the scientific case, we have two competing physical theories. If one of these theories is true, then the other one is not only false, but physically impossible. If Bohm's theory is true then, for example, it is physically impossible for particles to have indeterminate positions, as required by the Copenhagen theory. This is typical of cases of competing scientific theories. Thus, Dretske's account would imply that the two hypotheses in the scientific

case are not both relevant alternatives; whichever theory is false is an irrelevant alternative, because it is not genuinely possible. And therefore, Dretske's theory really would license the conclusion that one could know the Copenhagen interpretation to be true (assuming that it *is* true) even though one has no reason to reject Bohm's interpretation.

6. A reformulation of the skeptic's argument and the direct realist's response

This casts doubt on the validity of Klein's and Dretske's replies. However, all we've done so far is to pump the intuition that there is something wrong with those replies to the skeptic. We haven't actually explained what is wrong with them. Klein's response, at least, seems to work against the skeptical argument as formulated, so we need to reexamine the skeptic's argument.

The problem is that, as we formulated and defended the Closure Principle, your having justification for the claim that you're not a brain in a vat would be a *result* of your having justification for the claim that you have a body. But what the skeptic wants to say is that your having justification for the claim that you're not a brain in a vat is a *precondition* on your having justification for the claim that you have a body—that you need to *first* be in a position to know you're not a brain in a vat *in order* to be justified in believing that you have a body. So the Closure Principle,

(1) If S is justified in believing P and P entails Q, then S is justified in believing Q,

doesn't do justice to the skeptic's motivating idea.

Of course, it would not be acceptable to merely substitute the following, logically stronger principle:

(5) If P entails Q, then a *precondition* on S's being justified in believing P is that S be justified in believing Q.

This principle has no intuitive plausibility. For one thing, it would entail that a precondition on being justified in believing P, for any P, is that one be justified in believing ~~P, and also a precondition on being justified in believing ~~P is that one be justified in believing P; so one could never be justified in believing anything. No, the skeptic needs to say something more specific about the relationship between the brain-in-a-vat hypothesis and the claim that I have a body than that one entails the negation of the other. The skeptic needs to formulate an epistemological principle weaker than the absurd principle (5) above but still entailing that ruling out the brain-in-a-vat hypothesis is a precondition on knowing I have a body. At the same time, we want this epistemological principle, whatever it is, to account for our intuitions about the courtroom case and the scientific case discussed above.

So here's what we're looking for: we want an epistemological principle that, first of all, shows why in the courtroom case, we cannot merely grant that the defense hypothesis of a police conspiracy may be true and still claim to know that S is guilty. It should at the same time explain why, presumably for the same

reason, we cannot merely grant that Bohm's interpretation of quantum mechanics may be correct and still claim to know that the Copenhagen interpretation is the right one. The Closure Principle, of course, would satisfy this desideratum. But second, we want the principle to explain why in the courtroom case, the defense attorney's argument does not beg the question, and in the scientific case, the physicist criticizing the Copenhagen interpretation is not begging the question either. This is where the Closure Principle falls short, because it does not tell us why the received view in these cases couldn't count as a source of justification for rejecting the rival hypotheses. Finally, we want our epistemological principle to explain why one *might think* the brain-in-a-vat argument was sound. We don't actually want to make the brain-in-a-vat argument out to *be* sound; in fact, it is a bonus if we can explain why the brain-in-a-vat argument is not sound, even though it might reasonably appear so.

Now consider the following epistemological principle, which I will call the "Preference Principle" (because it concerns the preference of one hypothesis over another):

(6) If E is any evidence and H_1 and H_2 are two incompatible explanations for E, then S is justified in believing H_1 on the basis of E only if S has an independent reason for rejecting H_2.

In this context, an "independent reason" means a reason distinct from H_1 and not justified, directly or indirectly, through H_1. So the idea is that when you're faced with two competing explanations of certain data, you can't accept the one explanation until you have first ruled out the other. One's reasons for rejecting H_2 might include *a priori* reasons, such as that H_2 is significantly less simple than H_1, as well as empirical reasons.

Notice how the Preference Principle is weaker than the principle (5) that we rejected above. (5) would require us to be able to rule out each logical contrary of H_1 (in the sense of having reason to accept its negation), in order to be justified in accepting H_1. Thus, for example, we would have to be able to rule out $(\sim H_1 \ \& \ Q)$, where Q is any arbitrary proposition, as a precondition on being justified in accepting H_1. But the Preference Principle doesn't demand this. It only concerns the alternative *explanations of the data*. If H_1 is an explanation of E, $(\sim H_1 \ \& \ Q)$ will not generally be an explanation of E. For instance, Newton's Theory of Gravity (along with background assumptions) is an explanation for the fact that things fall to the ground when dropped. But the proposition, "Newton's Theory of Gravity is false and my socks are white" is not an explanation for the fact that things fall to the ground when dropped. So in order to accept the Theory of Gravity, we are not required to have an independent reason for rejecting, "the Theory of Gravity is false and my socks are white." This is fortunate, since the only reason I in fact have for rejecting that proposition is the Theory of Gravity.

The Preference Principle seems plausible intuitively, and it satisfies our desiderata. In the courtroom case, the hypothesis that S is guilty and the hypothesis that S was framed by the police are two competing explanations for the fact that S's blood was found at the crime scene, so we cannot accept that S is guilty on the basis of that evidence unless we rule out the other hypothesis.[15] Also, relying on the Preference Principle, the defense attorney is not open to a charge of begging

the question. To assert that we have *no reason* for rejecting the defense hypothesis may require begging the question, because in order to establish that we have no such reason, one must first establish that we don't know S is guilty. However, in applying the Preference Principle, the defense attorney only need assert that we have no *independent* reason for rejecting the defense hypothesis, i.e., no reason that is independent of the claim that S is guilty. And to argue that we have no reason *independent of the claim that S is guilty* for rejecting the defense hypothesis clearly does not require one to first establish that we don't know that S is guilty. Similarly, for the scientific case, we have two competing hypotheses, so according to the Preference Principle we must rule out Bohm's theory before we can accept the Copenhagen theory, and we must do so on grounds independent of the Copenhagen theory.

Now when we turn to the brain-in-a-vat argument, we can see why the argument would appear to be sound and non-question-begging—*if* one accepts one of the assumptions of indirect realism. If one accepts that beliefs about the external world are hypotheses for which the evidence is that we have certain sorts of sensory experiences, then the Preference Principle comes into play. Frank Jackson states this view particularly clearly:

> Our beliefs about objects, *all* of them (including the ones about causal links between sense-data and objects), form a theory, "the theory of the external world," which is then justified by its explanatory and predictive power with respect to our sense-data.[16]

Our ordinary, common sense beliefs about the external world, on the one hand, and the brain-in-a-vat hypothesis, on the other hand, are then two competing explanations for the same data. Therefore, just as in the courtroom case and the scientific case, we must rule out the brain-in-a-vat hypothesis in order to be justified in accepting our common sense beliefs about the external world on the basis of that data. So the indirect realist is faced with the responsibility of refuting the brain-in-a-vat hypothesis.

On the other hand, we can also see why we need not accept the brain-in-a-vat argument with its skeptical conclusion—if we adopt a direct realist account of perception. For the direct realist, perceptual beliefs about the external world are foundational; they are not hypotheses posited to explain anything. *Some* beliefs about the external world are hypotheses posited to explain evidence, such as atomic theory or electromagnetic theory; but immediate perceptual beliefs such as "Here is a red, round thing" are not. So the direct realist is in a position to make a principled distinction between, on the one hand, the courtroom case or the scientific case, where alternative hypotheses *do* need to be ruled out; and, on the other hand, the case of our ordinary perceptual beliefs. In the courtroom case and the scientific case, we really do have hypotheses posited to explain certain data, and as a result, the justification of a particular hypothesis depends upon a claim of superiority for that hypothesis over the alternative explanations.

Furthermore, the direct realist is in a position to explain simply how I know I'm not a brain in a vat. When I look at my two hands, for example, I know directly that I have two hands. It follows from this that I am not a brain in a vat. Notice that what is a question-begging argument for the indirect realist is not

question-begging for the direct realist. For the indirect realist, the argument just proposed is circular, because I have to start with the mere fact that I have certain sorts of experiences. From there, I don't have any way of getting to the claim that I have two hands except by ruling out the alternative explanations of those experiences. So I can't use the fact that I have two hands to rule out skeptical alternatives. But the argument is *not* circular as proposed by the direct realist, because I'm allowed to *start* from the claim that I have two hands. I'm not required to give an argument for that, so in particular I do not have to give an argument for it that presupposes the conclusion that I'm not a brain in a vat. The conclusion that I'm not a brain in a vat can be justified by a linear argument starting from foundational propositions.

7. Two objections

Objection #1

The direct realist line gets us out of the skeptical problem. But does it perhaps get us too much? There are some circumstances in which we genuinely need to consider alternative "hypotheses" to our perceptual judgements. We do not want our epistemological theory to rule out all such circumstances automatically. We don't want our response to the brain-in-a-vat argument to turn into a recipe for dogmatism with respect to perceptual beliefs.

Here is an example of the sort of circumstance I have in mind. Suppose I am driving late at night. There's a stone wall running along the side of the road. And suppose I seem to see a ghostly white figure at the side of the road walk through the stone wall, at a place where there is no opening. Now I can consider a few different hypotheses. One possibility is that I just saw a ghost walk through a wall. Another possibility is that there was actually an opening in the wall that I somehow did not see, and I saw a person who was walking through it. And a third possibility—the "skeptical" hypothesis if you like—is that there was neither person nor ghost there at all, and I merely hallucinated it. In this circumstance, it seems that I should weight the advantages and disadvantages of the possible explanations for my experience, as the Preference Principle would suggest. In fact the rational conclusion seems to be the "skeptical" one.

But wouldn't my direct realism enable me to resist this, just as it enables me to resist the brain-in-a-vat argument? Suppose I say that I have foundational knowledge that the white figure just walked through the wall, and since this entails that I did not *merely* hallucinate the figure, I can easily rule out that skeptical hypothesis. Isn't this comparable to claiming that since I have foundational knowledge that I have two hands, I can rule out the brain-in-a-vat hypothesis?

The key to unraveling this objection is the notion of *prima facie* justification. The direct realist need not—and should not—hold that perceptual beliefs have a kind of justification that is immune from countervailing considerations. He should hold that the justification attaching to immediate perceptual beliefs is, while foundational, nevertheless defeasible justification. The idea here is similar to the legal concept of presumption: perceptual beliefs may be presumed true unless and until contrary evidence appears. As long as there are no special grounds for doubting a given perceptual belief, it retains its status as justified, but

when other, justified or *prima facie* justified beliefs start disconfirming it, the presumption in favor of the perceptual belief can be defeated and the perceptual belief can wind up unjustified.[17]

This is the case in the example just described. I have a certain degree of *prima facie* justification for thinking the ghostly figure just walked through the wall. As I might say, *if I didn't know better*, I would naturally (and reasonably) assume that that is what happened. However, I have a large body of background knowledge, which indicates among other things that people generally can't walk through walls and that ghosts probably don't exist, and this defeats my justification for thinking the figure walked through the wall. It seems clear that this must be the right analysis of the case—as opposed to the view that we always need to rule out the possibility of hallucination before accepting perceptual beliefs—because in cases in which there is no evidence either for or against the hypothesis of hallucination (e.g., if I had merely seen a rabbit sitting by the side of the road), our default assumption is that things are the way they appear.

Now the brain-in-a-vat hypothesis is different. There are no grounds for suspecting that I'm a brain in a vat, in the way that there *are* grounds for suspecting that my seeming ghost sighting is a hallucination. So the presumption in favor of my perceptual belief that I have two hands, for example, remains undefeated, and this belief is therefore available for constructing an argument against the brain-in-a-vat hypothesis.

Objection #2

Does my response to skepticism merely beg the question? The skeptic's position is that we are not justified in believing any contingent propositions about the external world. I have responded to the skeptic by putting forward a direct realist account of perceptual knowledge. In asserting direct realism, I am asserting that we have a certain kind of justification for certain propositions about the external world. So at least part of my direct realist thesis is simply the negation of the skeptical thesis—namely, that we *are* justified in believing some propositions about the external world. Doesn't this mean that my response merely begs the question against skepticism? Obviously, the skeptic will just immediately reject direct realism. How does my asserting direct realism constitute any more of a response to the skeptic's position than just saying, "Skepticism is false"?

In answer to this objection, we need to distinguish two senses in which one might give a "response to the skeptic." One way to respond to the skeptic would be to give a positive argument, addressed to the skeptic, to show that we *do* have knowledge of the external world. This we might term an *aggressive* response to the skeptic. As the above objection shows, I have not given an adequate response of this kind. In fact, I do not believe it is possible to give a non-question-begging, positive argument against skepticism.

However, another sense in which one might be said to respond to the skeptic is this: one might confront an argument produced by the skeptic that tries to show that we *don't* have knowledge of the external world, and show how our common sense beliefs can be defended in the face of that argument. That is, one might demonstrate how the skeptical argument fails to give us a good reason for think-

ing that we don't know about the external world. We can call this sort of response a *defensive* response to the skeptic.

That is the sort of response I have provided. Given this aim, my asserting direct realism is a legitimate and non-question-begging move. I do not put forward direct realism as a premise from which to prove, positively, that we can know about the external world. That would certainly beg the question. Rather, I argue that the skeptic has only refuted one possible account of our knowledge of the external world, namely, indirect realism. I put forward direct realism by way of showing that there is an alternative account of our knowledge of the external world that is not damaged by the skeptical argument. The point is that if we take the direct realist line, then the skeptic hasn't given us any non-question-begging grounds for changing our position. The skeptic has merely assumed that we will take the indirect realist line.

8. Conclusion

Now let's conclude with a review of what I have and haven't done. I have proposed an epistemological form of direct realism according to which perception gives us a kind of justification for certain beliefs about the external world that is

(a) foundational, in the sense that the perceptual beliefs are not based on any other beliefs, but
(b) defeasible, in the sense that countervailing evidence can rationally require us to revise the perceptual beliefs.

I have not sought to elaborate and argue for this theory in any detail. Rather, I have focused on demonstrating one important advantage that a theory of this kind has over indirect realism, an advantage that has hitherto been overlooked by direct and indirect realists alike. I have shown that, whereas the indirect realist has an obligation of refuting the brain-in-a-vat hypothesis on grounds independent of our common sense beliefs about the external world, the direct realist can easily refute the brain-in-a-vat hypothesis on the basis of his beliefs about the external world. We saw that this does not involve the epistemological direct realist in circular reasoning, since he is able to construct a valid deductive argument starting only from foundational propositions. Finally, we have seen that the direct realist is able to handle the following cases in the intuitively acceptable manner:

(i) The courtroom case, in which we imagine that a jury member arguing for a conviction rejects the defense attorney's alternative explanation of the evidence, on the ground that the defendant is guilty,
(ii) the scientific case, in which a physicist rejects Bohm's interpretation of quantum mechanics on the sole ground that it contradicts the received interpretation, and
(iii) the case of the ghost sighting, where we imagine that I argue that my apparent ghost sighting could not have been a hallucination, since the figure in white really did walk through the wall.

The direct realist can distinguish each of the above examples of bad reasoning from his own reasoning against the skeptic. The first two cases are disanalogous because they both involve hypotheses inferred from evidence, whereas perceptual beliefs are not hypotheses inferred from evidence. The third case is disanalogous because it is a case in which specific reasons for doubting what I appear to have seen have defeated the initial justification for my perceptual belief, but there are no such defeaters for perceptual beliefs *in general*.

Thus, we've revealed a new way in which direct realism comes to our aid in fending off skepticism. I haven't shown that the indirect realist is inevitably committed to skepticism, since the indirect realist might still come up with a way to argue against the brain-in-a-vat hypothesis on *a priori* grounds.[18] The direct realist's advantage is simply that he doesn't need to go down that road—he doesn't have to play the skeptic's game to begin with. Moreover, the direct realist does not make out our knowledge of the external world to be contingent on any abstract, recherché reasoning of which only a small percentage of people in the world are aware. Refuting the brain-in-a-vat hypothesis is not a precondition on having knowledge of the external world, and such knowledge is well within the reach of all normal human beings.[19]

Notes

1 See Thomas Reid, *Inquiry and Essays* (Indianapolis, Ind.: Hackett, 1983), pp. 10–11 and throughout.

2 Compare Richard Fumerton's definition of "epistemological naive realism" in *Metaphysical and Epistemological Problems of Perception* (Lincoln, Nebr.: University of Nebraska Press, 1985), p. 73.

3 This principle will need some qualifications to protect it from easy counter-examples. For instance, if Q is a necessary truth, then P entails Q no matter what Q is, yet it is not apparently the case that, by knowing that the sky is blue, I am justified in believing Gödel's Incompleteness Theorem. We can avoid this problem by restricting the application of the principle to contingent propositions. We can further qualify the principle by restricting it to cases in which S is able to see that P entails Q. (See Peter Klein, "Skepticism, and Closure: Why the Evil Genius Argument Fails," *Philosophical Topics* 23 (1995): 213–36, p. 215.) These qualifications don't affect the skeptical argument or the responses to it discussed below.

4 Compare Peter Klein's remarks, *op. cit.*, p. 216.

5 See, for example, John McDowell, *Mind and World* (Cambridge, Mass.: Harvard University Press, 1996), pp. 111–13; Jonathan Dancy, "Arguments from Illusion," *The Philosophical Quarterly* 45 (1995): 421–38; and John Hyman, "The Causal Theory of Perception," *The Philosophical Quarterly* 42 (1992): 277–96.

6 *Op. cit.*, p. 294.

7 See, e.g., Robert Audi, *Belief, Justification, and Knowledge* (Belmont, Calif.: Wadsworth, 1988), p. 77; and Fred Dretske, "Epistemic Operators," *Journal of Philosophy* 67 (1970): 1007–23. But note that Dretske is discussing the Closure Principle for *knowledge* rather than for *justification*.

8 Dretske, *op. cit.*, p. 1016 (emphasis Dretske's).

9 Klein, *op. cit.*, pp. 218–22.