

7

How Do We Know that We're Not Brains in Vats?

Toward a Picture of Knowledge

7.1. Two Types of Explanation-Based Answers to Skepticism and the Problem with Giving Only the Positive Explanation

One way to present an anti-skeptical position is to advance an account of what knowledge is on which we do know what the skeptic denies we know. This could perhaps be a full-blown theory of knowledge—though, hopefully, a good enough picture would do.¹ One could then also put forward an account of what the relevant facts of our situation are, such that, given the facts posited, we satisfy the proposed requirements for knowledge, and so come out knowing at least much of what we take ourselves to know. (But as the account of our place in the world likely won't have to go beyond the claim that our factual situation is pretty much what we commonly take it to be, it can often just be left unsaid.) One could then apply such an account to the skeptic's argument, and the account may tell us where, and not just that, the argument goes wrong.

When applied to AI, which is so thoroughly driven by thoughts about knowledge (or the lack thereof), we will check our account's ruling on whether we know that skeptical hypotheses are false, in addition to its ruling that we do know the various Os we take ourselves to know, and we should be able to locate where, on our account, that argument goes wrong. On some accounts: "Oh, that argument misfires at its second premise—that's the one that's wrong on my view." We can call this the "positive explanatory approach" to skepticism (or the approach on which one provides a "positive explanation"). It is at least roughly the way that Nozick engages skepticism, utilizing what should be counted as a theory, as opposed to just a picture,

¹ The idea then would be that since the picture gives conditions for knowledge that are *approximately* correct, our satisfying the conditions it posits gives us good, even if not conclusive, reasons for concluding that we do know what the skeptic denies we know.

of knowledge. What he hopes to achieve in this way is an explanation, in light of the skeptic's argument, of how knowledge is possible (Nozick 1981: 8–18).

More generally, positive explanatory responses would go like this:

The skeptic argues that premise 1; premise 2; premise 3;...and so C. We don't know that O. But what if knowledge required just so-and-so, and our situation were such-and-such? Then (say) premise 4 would be false and the skeptic's argument would fail, and, what's more, we would know that O. On my account, that's just what happens. And so that is my account of how, in the face of the skeptic's argument, we [might] know that O.

The bracketed "might" toward the end of the response can be included, or not, to taste, depending on how assertive the anti-skeptic is feeling, the less assertive following Nozick in presenting theirs as an account of how, in the face of the skeptical argument, knowledge is *possible*. (I've always taken it that a mere possibility is claimed largely to acknowledge the potentially disputed nature of the anti-skeptic's picture of our factual situation.)

To the extent that one's account is plausible, this could I suppose be counted as explaining, in the face of the skeptical argument in question, how it is that we [might] know what we take ourselves to know.

I have followed Nozick in some important ways in how I approach skepticism. Like Nozick and others before me and around me, as we have seen, I eschew attempts at non-question-begging refutations of skepticism or proofs that the skeptic is wrong. And like Nozick, what I instead offer is an *explanation*, based on ideas about what is involved in knowledge. But there has also been a crucial difference. Where Nozick seeks to explain *how it is that we know*, for dialectical reasons we have seen, my focus instead has been on explaining *why it can seem that we don't know*: To explain away the intuitive appeal of the skeptic's case. If it's premise 4, say, of the skeptic's argument that we're going to deny, and that premise can seem quite plausible, I want to account for why, despite being false, that premise is tempting. We can call this the "negative explanatory approach" to skepticism. (And as we've seen, an important element of Nozick's view is central to my endeavors. So I have found part of Nozick's treatment of knowledge and skepticism to be better suited to the negative explanatory task that I have adopted than to the positive one he himself attempted.)

To quickly review Nozick's own treatment of AI and my complaints about it: Nozick denies AI's second premise. He admits the plausibility of what he's denying. In fact, he likens the closure principle on which this premise could be based, in terms of its "intuitive appeal," to a steamroller.² But following the positive explanatory methodology I ascribed to him above, he denies it because his account of knowledge

² "Uncovering difficulties in the details of particular formulations of [closure principle] P will not weaken the principle's intuitive appeal; such quibbling will seem at best like a wasp attacking a steamroller, at worst like an effort in bad faith to avoid being pulled along by the skeptic's argument" (Nozick 1981: 206).

rules that it is false.³ So, to now ask the question that vexes negative explainers like me: What then leads us so very badly astray about closure and AI's second premise? It's here that Nozick offers nothing, simply leaving that explanatory task to "further exploration."⁴

As I complained in SSP (Section 1.9), this leaves us with little reason to follow Nozick in denying just the piece of the puzzle he chooses to deny. It is worth pausing to clarify my complaint. The fault may be largely mine (perhaps due to my use of the strong word, "abominable"), but I often find philosophers thinking my "abominable conjunction charge" (as I've heard it called) is meant to just blow Nozick's theory out of the water by itself, or constitute some kind of refutation of Nozick's approach. (And some seem to think it succeeds at that aim.) But really, it is only intended to secure a place at the table, as a piece of the AI puzzle that we should seek to do justice to, for the comparative matter embodied in AI's second premise.⁵ The real objection against Nozick, then, isn't just that he's "subject to the 'abominable conjunction.'" That *couldn't* be the real objection here, because that was just intended as the kind of surface problem that any approach to the puzzle will have. The game is to explain such surface problems away. The real objection is that Nozick doesn't do anything to explain away the particular surface problem that he chooses to adopt by taking the view that he does. Once we see what the game is, or should be, Nozick doesn't even get into the game, by my lights.

The positive explainer might be leaning on the supposed success of his account of knowledge, especially if he is putting it forward, not just as a promising picture, but as a successful full-blown theory. I think that is Nozick's reason for why we should follow him. Analyzing knowledge has of course proven an extremely difficult game (some of the history of which is nicely presented in Shope 1983). If someone actually won that game, that would be a victory to which plausibly many spoils would be due, perhaps including following the theory's verdict on the skeptical puzzle: "Well, I really wish we had some insight into what made this such a puzzle to begin with (and hopefully 'further exploration' will reveal this), but, for crying out loud, we've finally reached the promised land and now have a winning theory of knowledge in our possession! We should accept its rulings on our puzzle. If it says it's AI's second

³ "Principle P is wrong, however, and not merely in detail. Knowledge is not closed under known logical implication. S knows that p when S has a true belief that p, and S wouldn't have a false belief that p (condition 3) and S would have a true belief that p (condition 4). Neither of these latter two conditions is closed under known logical implication" (Nozick 1981: 206).

⁴ "Thus, if our notion of knowledge was as strong as we naturally tend to think (namely, closed under known logical implication) then the skeptic would be right. (But why do we naturally think this? Further exploration and explanation is needed of the intuitive roots of the natural assumption that knowledge is closed under known logical implication)" (Nozick 1981: 242).

⁵ It is then very much like my attempt, in Section 5.1, to get the proper respect for the very similar comparative matter in play in the Harman lottery puzzle.

premise that's wrong, that's good enough for me.”⁶ It's in light of the result that Nozick's proposed analysis of knowledge turned out to have the same kind of problems that its many predecessors had, and so did not so stand as a victor to whom such spoils were due, that in SSP (see the second-to-last paragraph of Section 1.9) I concluded that Nozick's failure to offer anything at all on the negative explanatory task left us with no reason to follow his solution to our puzzle. But I now wonder about the coherence of wondering what we should have done if Nozick's theory had instead produced a victorious analysis of knowledge that then gave an implausible ruling on the comparative matter embodied in AI's second premise. I was able to easily suppose that because I was imagining the analysis game being played as it often is in epistemology, where accounts are judged by their intuitive enough handling of individual judgments of whether characters know in particular examples. But why shouldn't analyses be similarly responsible to deliver intuitively non-jarring results about comparative matters, as well? If they were, then Nozick's counter-intuitive pronouncement against AI's second premise would count against his theory being victorious in the first place. But I suppose the thought might then be that the theory had done so well at other applications that we should stick by it even when it runs into a little trouble. At any rate, I take it that even when the analysis game is narrowly construed, Nozick's theory has not passed through the valley of the shadows of counterexamples unscathed, victorious, and ready to claim its spoils.

7.2. Combining the Positive and Negative Explanatory Approaches to Skepticism: Toward a Not-so-Partial Picture of Knowledge

So, perhaps in (over)reaction to what I saw as the deficiency of Nozick's mode of engaging skepticism, I have been *all about* explaining how we get misled in our thinking about AI.

Given my negative explanatory aim, I haven't needed even so much as a "picture" of knowledge, but could make do with what we can just call a "partial picture." My use of "partial picture" of course is meant to convey that such an account can be even less fully worked out than is a (not-so-partial) "picture," but more specifically, it is meant to do so by pointing to the different dialectical aim it is used to achieve. I am not trying to motivate acceptance of our knowing that O by means of an account of knowledge that is plausibly close enough to being complete and correct that our satisfying all the conditions it posits (given my view of our factual situation) gives us good reason to accept that we know. Rather, following the Moorean methodology we

⁶ I imagine that some might instead be inclined in that case to only give Nozick his ruling on AI in the sense of accepting that he's identified the argument's problem *if our factual situation is as we suppose*, but then to insist on the skeptic's behalf that we cannot in this battle so suppose we're right about our place in the world.

looked at in Chapter 2, I have not been seeking *any* further positive basis for accepting that we know that O. It should already be seen as having the same standing as the skeptic's premises. What we really need, and should be seeking, is an account of why AI has the skeptical persuasive power that it has that is (good and) better than the skeptic's rival attempt to explain away our conviction that we know that O.⁷ So I end up focused on explaining, as I've put it here, why, when faced with AI, it can seem that we don't know, rather than on explaining how it is that we do know. And for the former, negative task, all we need from an account of knowledge is enough to tell us what at-least-apparent ingredient of knowledge the skeptical argument can make it seem to us to be missing, which it can do even while we are allowing that there might be other crucial elements of knowledge that our account does not even touch on.

It was in that spirit that in SSP I adopted an account of what knowledge is. Its purpose was not to provide a complete, or even nearly complete, picture of knowledge (one close enough to being correct and complete that our satisfying the conditions it posits gives us good reason to think we know), but rather, together with my contextualist semantics (now thoroughly defended in volume 1, DeRose 2009), to achieve the negative explanatory goal.

But of course, it would be better to give a picture (and not just a partial picture) of what knowledge is, and provide both the positive and the negative explanations. After all, as I wrote in Chapter 2, following Moorean methodology, one of the pay-offs of finding a solution to the AI puzzle was supposed to be that doing so would give us guidance toward an account of knowledge. We can then present a total package: Here, at least roughly (this will still be just a picture, not a theory), is what knowledge is; here is how, in the face of AI, we know that O; and here's why the skeptical argument can tempt us to think that we don't. How nice!

So in what follows, I will seek to develop the account of knowledge that I use in SSP into, or at least toward being, a "picture," of knowledge, as opposed to just the "partial picture" that it is in SSP. I will be focused on developing the aspects of the account most relevant to its application to skepticism, and especially on the aspects of that application that seem most open to challenge. And I will also seek to be explicit about the ways that the emerging picture remains imprecise.

7.3. Pictures, Theories, and Examples

On the double-safety account of it that I used in SSP, knowledge is a matter of having a true belief that *p*, where one's belief as to whether or not *p* is the case could not

⁷ Here we skip over the complications produced by the contextualist nature of my solution. As we've seen, in an important way, mine is a skeptical solution, in that on it we don't know that O by the standards that the presentation of AI has some tendency to put into place. Having here introduced that complication, I can say that since I contend specifically with *bold* skepticism, what I need is a better account of how the puzzle arises than can be given from that bold skeptical standpoint.

easily have been wrong—or, put in terms of possible worlds, of having a true belief that *p* is the case, where there are no possible worlds too close to (resembling too closely) the actual world where one's belief as to whether or not *p* is the case fails to match the fact of the matter. But there are different ways of measuring the ease with which one's belief as to whether or not something is so could have been wrong, and of measuring the “closeness” of possible worlds to the actual world. It is in the lack of precision in my use of those terms, more than anything else, that my account of knowledge is a picture, at most, rather than a theory.

Of course, the precision of one's use of such terms is a matter of degree, and, my Kripke-inspired use of “picture” and “theory” being itself also very imprecise, it is at this point quite unclear what I mean by my (rather suspicious, in a way reminiscent of the suspiciousness of Kripke's own use of the terminology) claim to be providing “only a picture” here, I know. What is mainly behind my characterization is my very picturesque attitude toward examples that test the account. “Theories” are supposed to issue *predictions* about such cases, so we can judge them by whether those predictions prove right or wrong, and it seems they are supposed to issue these predictions based on one's offhand understanding of the key terms used in the “theories.” Critics who put forward examples about which they think the theories they're criticizing issue incorrect predictions think of themselves as producing *counterexamples* which refute the theories in question. That's how the counterexample game is played. But when they try that on me, they will find me, as in Chapter 6, instead taking their examples as showing how the relevant notion of closeness or ease is to be understood, or as suggesting modifications in the formulation of my account, and so as providing opportunities to further sharpen the picture of knowledge that is emerging, rather than as a refutation of the account. (Indeed, strangely, they sometimes find that I have already sharpened the picture in a way that handles their proposed counterexample, as we saw in Chapter 6.) It seems like cheating, I know. But I'm not cheating at their game; I'm playing a different (and I think better) game.

My attitude here is very much of a piece with, and is likely largely inspired by, the attitude toward alleged counterexamples displayed by David Lewis in his “Counterfactual Dependence and Time's Arrow” (1979b)—though I don't rely on Lewis's particular application of this methodology being sound. Lewis is there defending this analysis of counterfactual conditionals:

Analysis 2. A counterfactual “If it were that *A*, then it would be that *C*” is (non-vacuously) true if and only if some (accessible) world where both *A* and *C* are true is more similar to our actual world, overall, than is any world where *A* is true but *C* is false. (1979b: 465)

Several critics had put forward, as counterexamples to Analysis 2, cases like one due to Kit Fine (and this is the one Lewis himself explicitly displays), in which we imagine the world being such that “If Nixon had pressed the button there would have been a nuclear holocaust” is true, though no actual nuclear holocaust ever occurs. In some

ways of imagining the scenario (in which the counterfactual remains intuitively true), it sure seems as if there are A-and-not-C worlds (e.g., worlds where Nixon presses the button, but a malfunction prevents the disaster) more similar to our imagined (not-A-and-not-C) actual world than are the closest A-and-C worlds. After all, the occurrence vs. non-occurrence of a nuclear holocaust seems to be a *very* big difference among worlds! But Lewis has issued this “word of warning”:

It is all too easy to make offhand similarity judgments and then assume that they will do for all purposes. But if we respect the extreme shiftiness and context-dependence of similarity, we will not set much store by offhand judgments. We will be prepared to distinguish between the similarity relations that guide our offhand explicit judgments and those that govern our counterfactuals in various contexts. (1979b: 466)

And in keeping with that warning, he responds to the no-nuclear-holocaust example as follows:

The presence or absence of a nuclear holocaust surely does contribute with overwhelming weight to some prominent similarity relations. (For instance, to one that governs the explicit judgment of similarity in the consequent of “If Nixon had pressed the button, the world would be very different.”) But the relation that governs the counterfactual may not be one of these. It may nevertheless be a relation of overall similarity—not because it is likely to guide our explicit judgments of similarity, but rather because it is a resultant, under some system of weights or priorities, of a multitude of relations of similarity in particular respects. (1979b: 467)

Lewis then uses the example in question, along with some others, to arrive at an account of what the relation of similarity that governs the counterfactuals in question must be, put in terms of the relative importance of various aspects similarity in that relation (1979b: 472).

In a somewhat similar vein, my attitude toward examples—which may be put forward with very unfriendly intent—is that they help us to see how the account should be understood. This does not mean that examples could not be a key component of a good case for abandoning a picture (including my picture of knowledge); just that that case will have to be different from a simple refutation-by-counterexample. The questions we should be focused on should involve explanatory power: Ask not whether the semantically flexible items (here, the counterfactuals and related judgments of “closeness” of possible worlds) that appear in the account are used so that they line up well with how you are inclined to use the relevant terms when making offhand judgments; ask rather whether the account, when it uses the terms in the way it points to, well explains what it seeks to account for.

7.4. Developing the Picture: Closeness, Restrictions, and the Coordination of Safety with Sensitivity

The work I've done on developing my picture in response to examples has been primarily focused on sharpening the notion of sensitivity, as in Chapter 6. This is in

no small part because that's where much of the criticism has been directed, but this has not been an arbitrary choice by the critics in question, for it's at sensitivity that the picture is most effectively tested. But my approach to proposed counterexamples was evident already in SSP, prominently in my treatment, in Section 1.6, of Nozick's grandmother example, where I suggest that a proposed counterexample to Nozick's theory of knowledge and, by extension, to my own insensitivity account of the power of AI's first premise, be handled by an explanation of the way to understand insensitivity conditionals in contexts of epistemic evaluation, rather than by a modification of the wording of the sensitivity condition itself. Later, in Section 1.7 of SSP, and then in Sections 6.6–6.7 and 6.9–6.13, I focused on examples in which it seemed best to instead modify the explicit formulation of sensitivity. But whether the account of sensitivity is developed by means of specifying how the counterfactuals that dominate its formulations are to be understood, or rather by means of modifications of the explicit formulation of sensitivity, the idea is to *coordinate* the resulting double-safety account of strength of epistemic position with my account of sensitivity, so that these developments in our understanding of sensitivity will be mirrored by a similarly developing account of (double-) safety/strength epistemic position—which latter (and, importantly, *not* the former) amounts to my picture of what knowledge is.

So, for example, my own preference is to respond to Nozick's grandmother example by means of an explanation of how the insensitivity conditional is to be understood, in terms of how the closeness relation that governs it should be construed (as in the last two paragraphs of Section 1.6 of SSP). And so, the idea here would be that double-safety now be understood as involving that same way of measuring the closeness of possible worlds. But if I were to instead follow Nozick more closely, building talk of the method of belief formation into the explicit formulation of the sensitivity conditional, that too would be mirrored by a coordinated sharpening of the double-safety account of knowledge, which would then construe knowledge that *p* as a matter of one's belief as to whether *p* is the case matching the fact of the matter in all the near-enough worlds *in which one reaches one's belief via the method of belief formation one uses in the actual world*.

However exactly the developments proceed, the idea is that on the resulting coordinated, overall picture, there are two related and coordinated but importantly distinct ways of evaluating beliefs: (1) in terms of the strength of our epistemic position with respect to them, given in terms of a double-safety account of that; and (2) in terms of whether they are sensitive. The resulting picture's acceptability turns on its explanatory power. How well does a contextualist double-safety account of knowledge, together with something like my Rule of Sensitivity, explain the epistemic evaluations we make, in connection with the AI puzzle and elsewhere? And, importantly, recalling especially the aspect of my defense of insensitivity accounts in Section 6.5, we should compare how well this account explains the relevant phenomena with alternative accounts.

The account brings together what seem to be two quite different types of impulses toward judging that subjects don't know. Often we judge subjects not to know

something because, though their belief is true, it very easily could have been false: That there are very nearby (in some epistemically relevant sense) worlds in which they get the matter wrong seems a pretty good explanation of why they don't know. But there is another range of cases, which have their own distinctive, sometimes more tentative or sometimes conflicted "feel," in which we are at least pushed toward judging that subjects don't know because their belief is insensitive: They would have believed as they do even if they had been wrong. This, I have argued, is why we can seem not to know that we're not BIVs, though the central cases of this type would concern whether we "know" far less remote "skeptical hypotheses" to be false. As with the first variety of at-least-apparent ignorance, this second variety can also be viewed in terms of our getting the matter in question wrong in some alternative scenario: I seem not to know that I'm not a BIV because I believe that I am not a BIV in the closest worlds in which I am one. But this second variety of apparent ignorance can seem to be very different from the first, in a way that's most apparent where it's applied to BIV-like cases, for the knowledge-killing worlds in play in the second kind of case are often the reverse of "very nearby": If any worlds are "remote" in the relevant sense, it would be worlds in which we're BIVs!

But hope for bringing together these two varieties of impulses toward judgments of no-knowledge begins to arise when we see the possibility, and then the viability, of contextualism, as in volume 1 (DeRose 2009). Maybe in more general terms the cloud that hangs over the true beliefs in question of both types is that there seem to be relevant worlds at which the subjects are wrong about what they believe? Sometimes that can happen because there are worlds that are (in the epistemically relevant sense) very nearby at which we're wrong: This will be our first variety of apparent ignorance. As these beliefs fail to meet just about any standard for knowledge, we should expect the appearance of ignorance to be particularly strong in such cases. But if the standard for what counts as a near-enough-to-be-relevant world varies with context, and if there are some kinds of belief that have some feature in virtue of which we will tend to count as relevant to their evaluation worlds in which we go wrong about the matter in question (and so apply to these beliefs epistemic standards that they do not meet), then we will (tend to) count as not knowing the things we so believe. Such is the case, on my picture, for insensitive beliefs.

On contextualist, safety-like accounts of knowledge, the usefulness of some device like the Rule of Sensitivity is apparent. Suppose you are wondering whether someone "knows" that *p* by some unusually high standards (which, on safety-like accounts, involve unusually large spheres of epistemically relevant worlds)—perhaps because it is relevant to some very high-stakes concern you face. There are a variety of ways you could indicate that you intend such elevated standards in your use of "know(s)." For instance, you could tell the people you're talking to just how much is at stake, and hope that they adjust their epistemic standards up to something appropriate to the practical situation being discussed, and/or you could utilize something like rules of accommodation, steadfastly denying the status of "knowledge" to beliefs that don't

meet the stringent standards you intend, and only allowing beliefs that meet those standards to pass as “knowledge.” But if your intended standards are so high that a certain quite remote possibility of error that you have in mind should be counted as relevant, it would be very helpful to actually mention the remote possibility you have in mind and thereby indicate relatively precisely just how high are the standards are that you intend: Yes, for instance, you really do want to count as relevant such remote possibilities as that the bank has changed its hours in the past two weeks, or that for the first time ever, the reason John’s hat was hanging on the hook outside his office door was not because he was in his office, but because he forgot it when he left for home the day before. Given something like a Rule of Sensitivity, you can ask whether the subject knows these fairly remote scenarios don’t obtain, or (depending on what information you have) you can say that they don’t know that *p* because they don’t know that the fairly remote scenarios don’t obtain. This will often be an effective means of indicating that you’re not looking for someone who is just going by the presence of John’s hat, despite what a terrifically reliable indicator of John’s presence that has proved to be, but rather for someone who, for instance, clearly remembers having herself seen John at the office. (Of course, sometimes interlocutors may not want to go along with your intended use of “knows,” and you may meet resistance—just as may happen with certain intended uses of “tall.” This may especially happen when you try to install extremely high standards, where extremely remote possibilities of error are counted as relevant.)

7.5. Single- or Double-Safety?

The basic idea of normal “safety” accounts of knowledge (which I will call “single-safety” accounts) is that knowing that *p* is the case is a matter of having a true belief that *p*, where it could not easily have been the case that one believed that *p* but *p* was false—or, when the view is put in terms of possible worlds, of having a true belief that *p* is the case, where there are no possible worlds too close to the actual world in which one believes that *p* but *p* is false. Though such a safety account could have underwritten my negative explanation of the skeptic’s appeal in SSP, I instead found myself reaching for what I have since come to call a “double-safety” account as the partial or “toy theory” (as it can be thought of) of knowledge I used for the task, on which, as I have already mentioned, knowledge is a matter of having a true belief that *p*, where one’s belief *as to whether or not p is the case* could not easily have been wrong—or, put in terms of possible worlds, of having a true belief that *p* is the case, where there are no possible worlds too close to the actual world where one’s belief *as to whether or not p is the case* fails to match the fact of the matter. The difference here is that single-safety is only disturbed by too-nearby worlds where one believes that *p* but *p* is false, while double-safety is disturbed both by that and by too-nearby worlds in which *p* is true but one disbelieves *p*.

Perhaps because safety accounts of knowledge were not as prominent as they have since become when I chose a partial account of knowledge to use in my explanation of AI's appeal, I did not put all that much thought (and really none that I can recall) into whether to instead use (what has come to be a normal) single-safety account.⁸ I just grabbed right for the double-safety idea because it seemed to me the natural way to put into possible worlds talk what I took to be a basic safety-from-error approach to knowledge.

I am still inclined to opt for a double-safety account of knowledge, but that preference is still largely just based on my general sense of what knowledge that *p* would seem to involve, namely, safety from error on the matter of *whether or not p* is the case. Since my application of the account of knowledge to skepticism does not depend on the choice between these varieties of safety, I won't try to resolve the matter here. I will briefly indicate why I don't accept certain reasons for favoring one version of safety over the other, though, since that discussion may help to convey the picturesque attitude I take toward my account of knowledge.

I have in the past (some years ago, but some years after writing SSP) appealed to double-safety's better handling of cases of necessary truths as a reason for preferring it over single-safety. We can, of course, believe but fail to know necessary truths. The problem for single-safety here is that there can be no nearby worlds in which we believe a necessary truth, but in which it isn't true, since, being necessary, there are no worlds in which it isn't true. But the double-safety of such a belief can be upset by the presence of nearby (in the relevant way) worlds in which one disbelieves the necessary truth in question (DeRose 2004b: 34). Thus, suppose Uncle Lyle tells you twenty things, intending mainly to induce false beliefs in you by lying. However, he throws one truth into the mix, just to make things interesting. And suppose you do believe all twenty of the things you are told, just on Uncle Lyle's say-so. You had no opinion, nor any good basis for an opinion, on any of them before Uncle Lyle told you about them. Here, you seem not to know that one true thing, despite its truth and your believing it, even if it is a somewhat complicated mathematical truth that could not have been false. Since there are no worlds at all in which the proposition is false, there are no nearby worlds in which you believe it but it is false, so a simple single-safety condition cannot account for this lack of knowledge. However, double-safety can: You can fail to know in this case because there are dangerously nearby worlds in which you have a false belief about *whether* that mathematical proposition is true, namely, worlds in which Uncle Lyle chose some different truth to throw into the

⁸ I would have been doing this choosing in the very early 1990s, while writing SSP. The third chapter of my dissertation (DeRose 1990: 131–242) served as a basis for SSP, but in the dissertation I just used comparative conditionals to get a fix on relative strength of epistemic position, without any underlying partial account of what knowledge is, in order to get the result that we have to be in a stronger epistemic position with respect to beliefs that skeptical hypotheses do not obtain, as compared with beliefs in more ordinary things, for beliefs in the former to be sensitive, before utilizing (1990: 215–16) a form of what in SSP I came to call the “Rule of Sensitivity.”

mix, and so falsely told you the negation of the mathematical truth, rather than the truth itself.

However, that kind of counterexamplifying basis for rejecting single-safety no longer seems to me a good reason for preferring double-safety. David Manley has proposed a way to refine single-safety accounts to deal with such problems, by having them require for knowledge that p that there be no nearby worlds in which one falsely believes p or some other, related, “counterpart” proposition (Manley 2007).⁹ I suspect there will be examples in reaction to which I will want to develop even my double-safety account in Manley-like directions, in which case, there is reason to think that, having made those developments myself, I will lose the above reason for choosing against a single-safety account.

In a very similar way, but on the other side of things, I reject an argument Sosa has given for preferring single-safety over double-safety. Referring to (what I have now come to call) my “double-safety” condition as my “strength requirement,” and simply using “safety” to designate what I am here calling his own “single-safety” view, Sosa writes:

DeRose’s strength requirement is unacceptable as a necessary condition for knowledge, or so I will now argue. If I see a large pelican alight on my garden lawn in plain view, I will know that there is a bird in my garden. And this is not affected by the fact that a small robin sits in the garden in its nest out of view. In such circumstances, there might very easily have been a bird in the garden without my believing it. If the pelican had not arrived just then while the robin was still in its nest out of view, all of which might easily have happened, then there would have been a bird in my garden without my believing it. So, while DeRose’s condition is perhaps a condition for a kind of competence on the question whether p , it fails as a condition for knowledge that p . The safety condition, however, unlike the strength condition, still seems safely a live possibility. (Sosa 2004b: 280)

While this kind of case points to directions in which the basic safety approach to knowledge can be sharpened, I do not think it provides any real basis for deciding between single- and double-safety accounts. To see why, consider how Sosa’s counterexample can be easily modified so that it targets single- just as well as it does double-safety. Just take Sosa’s example, get rid of Sosa’s robin (so that now the only actual bird in the whole garden is the pelican), move the nest into view, and replace the robin in the nest with a weirdly shaped leaf that’s moving slightly in the breeze so

⁹ In email correspondence with Graeme Forbes, written during the time I was writing SSP and he was writing (Forbes 1995), I made somewhat similar suggestions for how to formulate a relevant alternatives theory of knowledge in light of problems similar to those that Manley is responding to on behalf of safety. I have lost that correspondence, but one of my examples is preserved in one of Forbes’s footnotes: “Keith DeRose suggested to me as an analogy to illustrate the point a case in which the subject is in a region filled with real but mobile barns which are switched around randomly at night. The subject’s belief ‘that’s a barn’ is knowledge because even in close possible situations where the externalist content of his belief is different because he is looking at a different barn, his belief is still true” (Forbes 1995: 215, n. 18). Manley develops his counterpart approach to safety both in response to such examples, and also to mathematical cases (the application needed here).

that, from our subject's position, it looks just like Sosa's robin, and would have fooled our subject into believing it was an actual bird if our subject's attention hadn't instead been grabbed by the very large and obviously real pelican that has alighted on her lawn. (Alternatively, you may suppose that the leaf *has* fooled our subject, who falsely believes that there are *two* birds in her garden.) Despite all the changes we have made, our new situation is still well viewed as a modification of Sosa's (as opposed to a new example, unrelated to Sosa's in epistemically relevant ways) in that the following specifications are the key to our new case and our evaluation of it: Make the pelican's presence every bit as chancy in our new case as it is in Sosa's own example, so that the worlds in which the pelican is not in the garden, but has alighted elsewhere, are every bit as nearby and threatening as they are in Sosa's case—it could have happened just as easily in our example as in Sosa's that the pelican was elsewhere. Still, as in Sosa's case, since the pelican did in fact end up in the garden, our subject, who clearly sees what is quite obviously a real pelican, knows that there is a bird in her garden. But now in our version of the case, the nearest worlds in which the pelican has alighted elsewhere are worlds in which our subject falsely believes there is a bird in her garden, and so makes the kind of misjudgment that ruins the single-safety, as well as the double-safety, of the belief she holds in her actual situation that there is a bird in her garden, should those worlds count as epistemically relevant.

I don't take my modification of Sosa's example to be a refutation of the single-safety (nor of double-safety) approach, because I think there are good ways to understand single-safety so that it gets the case right. (And readers of Chapter 6 should have some ideas of how this could be done.) But the best ways of getting Sosa's account of knowledge out of the trouble that the modified case seems to cause it also gets double-safety out of the trouble that Sosa's original example seemed to get it in. That's why I don't think Sosa's case can provide any good grounds for choosing single- over double-safety. This is just the kind of case that safety theorists of both stripes will have to deal with by means of the kinds of developments/sharpenings/modifications we have been discussing. There's no point in trying to wring an advantage of one over the other out of such cases.

7.6. Pointed Questions and Challenges Facing Contextualist Mooreans

In this chapter, what I am focused on is meeting the challenges my picture faces in its application to the skeptical argument. The particular challenges that I face in positively explaining in the face of AI how we know the Os we take ourselves to know are very much shaped by the fact that mine is a contextualist Moorean response to AI—and are challenges I share with other contextualist Mooreans.

Recall that "Mooreans" in our currently relevant sense (substantive, rather than methodological, "Mooreans") are those who hold that we do know that skeptical

hypotheses are false. Those most sensitive to the appeal of skepticism often raise the question of how we can know that we're not BIVs against Moorean responses in a pointed way: "How can we *possibly* know that we're not BIVs?" Answering this is a key challenge to (substantive) Mooreans.

Contextualist Mooreans (see Section 3.1) hold that we know that skeptical hypotheses are false *by ordinary standards for knowledge*, or, as we're here abbreviating it, that we know_o that ~H. As we noted in Section 3.1, this claim of contextualist Mooreans that we do have "regular," even if not "super-high-octane," knowledge of the fact that we're not BIVs exposes us to a version of the pointed question, since we can still be asked, perhaps quite pointedly: How can we possibly know, *even by just ordinary standards for knowledge*, that we're not BIVs?

So for a contextualist Moorean like me, the pressing challenge in positively explaining *in the face of AI* how we know_o that O is explaining how we know, even by moderate standards, that skeptical hypotheses like the BIV hypothesis are false. This is a pressing challenge first, and more intuitively, because of the impression that many can have that there is *no* good sense in which we could know such a thing—or at least that, even if there might be some especially low but allowable standards that are met by our beliefs that we're not BIVs,¹⁰ we don't know it by most ordinary standards. And second, after a bit of theory, the accounts we contextualist Mooreans have given of our alleged "knowledge" of such esoteric facts strike the skeptical as making that knowledge_o out to be problematically "empty" or "vacuous," and suspiciously a priori. (It also strikes some critics as suspiciously unspeakable, but this more language-based worry is answered in Section 4.16.)

Another problem my position can seem to have is that my account of knowledge can seem too radically externalist, leaving no room for a suitably internalist notion of justification as being involved in knowledge. As I answered in Section 4.13 (but now using the terminology just introduced here), this worry has been exacerbated by missing the fact that I was only engaged in (what we are now calling) the negative explanatory task in my earlier writings that were being targeted. Still, critics might well wonder how some suitably "internalist" account of justification can smoothly fit into the accounts of knowledge proffered by me and other contextualist Mooreans. And indeed, it was mainly because I had said nothing about epistemic justification that I took myself in SSP to be offering only a "partial picture"—an account intended only for use in the negative explanatory task needed to defeat the skeptic. It's only after some account of epistemic justification is added (or good answers are given for why knowledge involves no such thing) that we can be said to have entered even the level of pictures.

And, of course, then, at the intersection of the above two worries stands the (perhaps pointed) question: Even by just moderate standards (which would

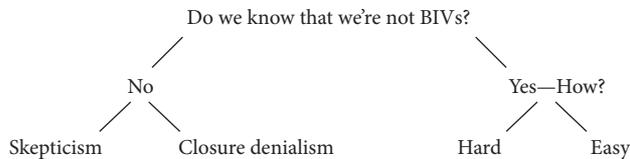
¹⁰ See DeRose (2009: 13–18) for a discussion of the possibility of especially low standards for knowledge.

presumably be the relevant ones for the kind of justification that might be an ingredient of knowledge_o), how can we possibly be justified in believing that we're not BIVs?

Beyond repeating that the resulting account will still be just a picture, three more bits of goal-post-setting clarifications, closely related to one another, are in order, before we proceed. First, we must realize from the start that this account of how we know that we're not BIVs can only get so plausible, for we are working against a fairly strong (or very strong, depending on the respondent and her current mood, see Sections 2.7–2.10) intuition to the effect that we do not know that we're not BIVs. Hopefully, the resulting picture will be plausible in the abstract as an account of what knowledge is, but when we apply it to \sim BIV, the application itself may have to be somewhat counter-intuitive, at least when viewed from certain angles. Second, when evaluating the success of the application of the resulting picture to the matter of whether we know we're not BIVs, we should remember that this positive account is backed by a negative explaining away of the intuition that we don't know that we're not BIVs. Third, of course, I *agree* that we don't know that we're not BIVs by some allowable standards for knowledge (in fact, the very ones that the discussion of the issue has some tendency to put into place), and I am accounting only for how we know that we're not BIVs *by ordinary standards for knowledge* (for how we know_o that we're not BIVs).

7.7. Keeping it Easy to Know_o that We're not Brains in Vats

Relative to our current concerns, it is handy to map out the various positions on the issues of whether and how we know that we're not BIVs as follows:



Contextualists' place on this chart is complicated by our contextualism. Relative to the exceedingly high standards that we think the skeptical argument has some tendency to put into place, we're in the far left, skeptical position: By those standards, not only do we not know that not-H, but we also don't know that O. But what we're interested in now is where we are on this chart with respect to knowledge by ordinary, moderate standards. And when it comes to such "knowledge," knowledge_o, contextualist Mooreans are by definition on the right-hand half of the chart, holding that by the ordinary standards at which we know that O, we also know that not-H.

And, recalling our discussion in Chapter 3, though it is not by definition, it is important to contextualist Mooreans to take the *far* right option and make it easy for

ordinary people to know by ordinary standards that they are not BIVs. We need not draw the distinction between “hard” and “easy” precisely, as we won't be trying to adjudicate any close calls. For our purposes, accounts of our knowledge that we're not BIVs make that knowledge “hard” to come by when they make it dependent on some fancy argument that the knower grasps to the conclusion that the skeptical hypothesis is false—an argument perhaps like Descartes', or Putnam's that starts from premises of a highly restricted and presumably exalted kind. We contextualist Mooreans seek to avoid such “heroism,” as I have called it (see again Chapter 3). To make our knowledge-by-ordinary-standards of not-H easy, non-heroes either (and this is my route, and seems to be that of contextualist Mooreans generally) make it out to be not dependent on any inference at all, but rather be had by some other means available easily to all, or else (and we will be taking a critical look at this alternative), if it is known by inference, by simple inferences from the likes of O itself.

In some important ways, I will follow the insightful lead of my fellow contextualist Moorean, Stewart Cohen, in wrestling with some of the issues that lie ahead in this chapter. In particular, I share some of Cohen's views about what a proper structure for our knowledge should look like, in terms of what can, and, more to the point, cannot, be properly based on what. But, though it's probably best counted a matter of packaging rather than substantive, one difference I have with Cohen that is important to note here (this seems to me important packaging), and that arises especially in his important paper “Basic Knowledge and the Problem of Easy Knowledge” (Cohen 2002), is that I don't think the problems we both have with our relevant opponents should be characterized in terms of those opponents making knowledge problematically *easy* to come by, as, for instance, Cohen does here: “the problem is that once we allow for basic knowledge, we can acquire reliability knowledge very easily—in fact, all too easily, from an intuitive perspective” (2002: 311). This is because I think the position that Cohen and I end up in also ends up making key bits of knowledge (for our current purposes, knowledge_o that radical skeptical hypotheses are false is what's most relevant) very easy for us to acquire—and that it's good and important for it to do so. Our complaints with the relevant opponents should be framed as a disagreement over how, and not whether, to make crucial bits of knowledge very easy to attain.

My talk of making our knowledge of \sim BIV easy to attain may cause some readers, wary of my contextualist ways, to think that I am here invoking unusually low standards for knowledge, and so only defending an *especially* low-grade “knowledge” of the likes of \sim BIV. (Sometimes in discussion, and in the literature, relatively low standards are labeled the “easy” standards.) So to avoid such misunderstandings, let me make clear that I don't mean anything like that by my use of “easy.” Of course, it's crucial to my position that we don't “know” that \sim BIV by the extremely inflated skeptical standards that the presentation of AI has some tendency to put into place, and so I will often write such things as that I am only claiming that we have knowledge by ordinary standards (knowledge_o) of the likes of \sim BIV—as opposed to our having knowledge by the AI skeptic's extreme standards. But my own position

is that we are in a very strong epistemic position indeed with respect to \sim BIV—strong enough to not just squeak over the hurdle of most ordinary standards for knowledge, but to match the strength of position we enjoy with respect to even such solid things as that we have hands (which also *easily* clears the bar set by ordinary standards, though it also doesn't meet the skeptic's extreme standards). My explicit claim in (Section 1.10 of) SSP is that we are in as strong an epistemic position with respect to \sim BIV as we are in with respect to the likes of *I have hands*. Of course, that surprising comparative fact can be due either to our being in a surprisingly strong epistemic position with respect to the former or, as the skeptic will urge, to our being in a surprisingly weak position with respect to the latter. The surprise, on my account, is how well-positioned we are with respect to our not being BIVs. This verdict is defended by means of an account of why it can misleadingly *seem* that we don't know that we're not BIVs at all, which can also serve (perhaps even more suitably) as an account for the more general phenomenon that we can tend to think our belief in \sim BIV constitutes very shaky knowledge at best (if it's knowledge at all). On my account, there is a conversational mechanism which has the effect of pushing our claims to “know” the likes of \sim BIV toward falsehood. It is the operation of this mechanism, not actual excessive shakiness, which explains why our “knowledge” that we're not BIVs can appear to be shaky at best—often to the extent that it seems not to be knowledge at all.

From my perspective, gathering evidence and constructing arguments is indeed one way, *but only one way among others*, to come to be in a strong(er) epistemic position with respect to propositions. Though success in that particular, difficult way often constitutes extremely impressive intellectual accomplishments on our part, those accomplishments are often episodes in our coming to meet some not-so-very-high standards with respect to the conclusions we draw. (Indeed, in line with my views expressed in Appendix C, I think that in many of the most impressive cases, we end up with conclusions that don't even meet quite ordinary standards for knowledge.) On my account, and on the type of account I think we should non-heroically seek, we are all¹¹ very well-positioned indeed with respect to our not being BIVs before we have any fancy arguments to that effect, and independently of any such arguments we might have.

7.8. Our Knowledge_o that We're not BIVs and the Charge of Vacuousness

So I seek to make our knowledge_o of \sim BIV out to be easy to come by. And the problem is that contextualist Moorean accounts have struck critics as making our

¹¹ Well, I trust everybody reading this: We can (and I will, in Section 7.20) dream up a character in weird circumstances who is in a poor epistemic position with respect to her not being a BIV.

alleged “knowledge” of the likes of our not being BIVs *too* easy to attain. Such criticism need not issue from skeptics or their friends; it can and in fact does come instead from closure denialists.

One worry along these lines that has been pressed by closure denialists—first by Mark Heller (1999: 207), and then, following Heller and expanding significantly on his work, more recently by Wesley Holliday (2015: 113–20)—is that the “knowledge” that we contextualist Mooreans claim that we have to the effect that we are not BIVs turns out to be problematically *vacuous*. Especially in Holliday’s hands, this charge is intertwined tightly with other worries he has about the “knowledge” we allege,¹² but it’s worth ripping it out from its connections with other objections to be answered on its own, getting to the associated worries later, since this particular objection should receive a response quite different from the responses to the related worries—namely, that the account of knowledge I have adopted turns out not to be subject to this precise charge (though also that it would not bother me if it were). The objection as I will quite narrowly understand it here is that the key component in our accounts of knowledge (beyond the standard conditions of true belief) is only vacuously satisfied by our beliefs that we’re not BIVs. Contextualist Mooreans have typically adopted contextualist versions of the relevant alternatives account of knowledge, on which the main ingredient which must be added to true belief in order to yield knowledge is that the believer rule out all the relevant alternatives to what she believes. And we then turn out to know_o that we’re not BIVs because, relative to standards *o*, there just are no relevant alternatives, ruled out or not, on which we are BIVs, and so it is the case, but only vacuously so, that there are no relevant alternatives on which we are BIVs that we have not ruled out.

Also hovering about, and especially relevant to my own contextualist solution to AI, are (fairly closely related) safety accounts of knowledge, on the normal version of which it also turns out that our knowledge_o of our not being BIVs seems vacuous. The basic idea of normal, or what I am calling “single,” “safety” accounts of knowledge, as I have noted earlier in this chapter, is that knowing that *p* is the case is a matter of having a true belief that *p*, where it could not easily have been the case that one believed that *p* but *p* was false—or, when the view is put in terms of possible worlds, of having a true belief that *p* is the case, where there are no possible worlds too close to the actual world in which one believes that *p* but *p* is false. As I have already mentioned, though such a safety account could have underwritten my negative explanation of the skeptic’s appeal in SSP, I instead found myself reaching for what I have since come to call a “double-safety” account, on which knowledge is a matter of having a true belief that *p*, where one’s belief *as to whether or not p is the case* could not easily have been wrong—or, put in terms of possible worlds, of having a true belief that *p* is the case, where there are no possible worlds too close to the

¹² Holliday seems particularly worried about our making this knowledge_o out to be a priori—which worry I will address at length in this chapter.

actual world where one's belief as to whether or not p is the case fails to match the fact of the matter. Again, the difference here is that while single-safety is only disturbed by too-nearby worlds where one believes that p but p is false, double-safety is disturbed both by that and by too-nearby worlds in which p is true but one disbelieves p .

Though, as I have noted, I did not put much thought into my choice, it turns out that my opting for a double-safety picture, instead of a normal, single-safety one, allows me to avoid the precise charge of vacuousness (as distinguished from nearby, associated worries), so far as I can see. It's when the safety account is put in terms of possible worlds that the specter of vacuity seems to arise: Your belief that you're not a BIV then satisfies the normal safety condition vacuously—there just are no too-nearby worlds in which you're a BIV, so it's vacuously true that there are no too-nearby worlds in which you're a BIV but think that you are not. By contrast, on my double-safety account, you know_o that you're not a BIV because (you correctly believe that you're not and) there are no too-nearby (by standards o) worlds in which your belief as to whether you are a BIV is wrong. With respect to your belief that you're not a BIV, then, you satisfy my double-safety condition (at ordinary strength) non-vacuously: Though they are all worlds at which you are not a BIV, there are indeed sufficiently nearby worlds at which my account requires you to have some correct belief, but you meet this requirement, quite non-vacuously (if somewhat monotonously), by having the right belief (that you're not a BIV) in those relevant worlds.¹³

7.9. What Is, and What Is Not, Important to Knowledge on the Basic Safety Approach: A Parable

I would like to cite its non-vacuous handling of our knowledge_o that we're not BIVs as an important advantage of my double-safety account of knowledge over normal, single-safety theories and relevant alternative accounts, but I don't really think it is such an advantage—for reasons pretty deeply embedded in the basic safety approach to knowledge. It's right in the name for the approach. Both single- and double-safety

¹³ The same is true of your belief in the conjunction that: You're not a BIV and you believe that you're not a BIV. With respect to your belief in that conjunction, you satisfy my double-safety condition (at moderate strength) non-vacuously: Though they are all worlds at which the conjunction is true, there are indeed sufficiently nearby worlds at which my account requires you to have some correct belief concerning that conjunction, but you meet this requirement, quite non-vacuously (if somewhat monotonously), by having the right belief (that the conjunction is true) in those relevant worlds.

This is in reply to Holliday's response to my claim that the double-safety account of our knowledge_o that we're not BIVs escapes the vacuousness objection (see Holliday 2015: 115). I should note that Holliday's response indicates that he is thinking of the objection of vacuousness as involving thoughts about whether this "knowledge" would represent a suitable "epistemic achievement," and, for reasons we will see in Section 7.9, I don't want to claim, as Holliday has the double-safety theorist claim, that this is such an "achievement." So I am construing the "vacuousness" objection more narrowly than Holliday appears to.

theorists are trying to articulate some kind of safety from error as the crucial condition for knowledge. (We differ over just which errors one must be safe from to know.) But if it's really safety from error that's important to whether one knows something, it wouldn't seem to matter whether that safety is "vacuous."

For a perhaps helpful analogy, consider this parable in which one's physical safety is well-measured by how close the nearest dangerous objects are. Suppose there are "killrocks" strewn about that are subject to emitting deadly energy. When a certain type of "triggering" event occurs (say, a pulse of otherwise harmless energy from a nearby star), killrocks will (convert that otherwise harmless energy into a sphere of death with very sharp boundaries and) kill all the plant and animal life sufficiently close to them, but how close is "sufficiently close" will vary by the strength of the triggering event: It could happen that they each kill everything within, say, ten feet of themselves, or, if the trigger is stronger, within thirty feet, or they could go off with deadly consequences for the life within any number of other possible radii. But suppose further that killrocks can be "neutralized" so that they won't emit deadly energy when a triggering event occurs. So, we are now supposing that one's safety can be measured by how close to one are the nearest *non-neutralized* killrocks. Now, suppose that Ann had found herself in a very dangerous situation, with non-neutralized killrocks very nearby to her: Almost any triggering event would have been the end of her. But suppose that through some very diligent, and sometimes very skillful and admirable, work, Ann has succeeded in neutralizing all the killrocks within forty feet of herself, and that consequently (given the properties of killrocks and how likely triggering events of various strengths are), she is now quite safe from death-by-killrock. Compare Ann's situation with that of Bill, who hasn't neutralized any killrocks at all, but finds himself in a quite safe situation nonetheless, because, though the killrocks that are nearest to him are not neutralized, there are no killrocks within forty feet of him. Bill's is a vacuous safety: There are no non-neutralized killrocks very close to him because there are no killrocks at all very close to him. Relatedly, Bill's safety, we may say, is "insensitive": Among the killrocks closest to him are non-neutralized ones. But though his safety is vacuous and insensitive in a way that Ann's isn't, and though he hasn't really done any "safety work," admirable or not, in virtue of which he is safe the way that Ann has, Bill is as safe from killrocks as Ann is. It would take the same kind of triggering event (a pulse so strong that it would cause the rocks to kill things more than forty feet from them) to kill Bill as to kill Ann, and, as we're supposing such events are very unlikely, our characters are both quite safe, and safe to the same high degree.

One could accept a safety condition for knowledge and still hold that an epistemic analogue of Ann is a better candidate for knowledge than is an analogue of Bill, though she is no more safe from error than he is, if one held that there was also some quite different important component of knowledge relative to which Ann does much better than Bill. Perhaps the most likely way to do that would be to include as an ingredient of knowledge some kind of justification on which an analogue of Ann, in

virtue of her praiseworthy, impressive, and effective intellectual efforts, scores more highly than does an analogue of the lucky, and perhaps lazy, Bill. But the basic safety approach runs too deeply in my thought to allow me to be happy with such a maneuver. Though I will accept a justification requirement for knowledge, it is not one that, with respect to a typical person's belief that they're not a BIV, upsets the ruling of basic safety approaches that an epistemic analogue of Bill is as good a candidate for knowledge as is an analogue of Ann. I not only think that safety from error is an important component of knowledge, but also that it is an important aspect of knowledge *as opposed to* these other (related) properties, which are not, nor of construals of justification which would align with them: Sensitivity, safety in virtue of having done some "epistemic work," safety that represents an "epistemic achievement,"¹⁴ non-vacuous safety. In light of that last item not being any part of what's important to knowledge, I don't take my account's non-vacuous handling of our knowledge, that we're not BIVs to be a good reason for choosing it over a normal, single-safety account.

But, more importantly, considering now all the items on the list, it's central to the (substantively) Moorean, safety approach to knowledge, at least as I see it, that lots of our knowledge just is extremely easy for us to possess. Our knowledge that does arise by hard "epistemic work" and that represents an "epistemic achievement" on our part rests on other knowledge we have that is basically just handed to us, and our knowledge, of the likes of our not being BIVs is the kind we possess easily.

7.10. Radical Skeptical Hypotheses, A Prioritism, and Split Cases AI

To a substantial degree, we contextualist Mooreans take the same approach to AI (our basic skeptical argument by skeptical hypothesis), whether the skeptical hypothesis there utilized is just a *fairly* remote one, like that yonder animals in the zebra cage at the zoo are cleverly painted mules, or an *extremely* remote, radical, global one, like that one is a BIV. In both cases, we say that you know that $\sim H$ by the same *relatively* low standards that you know that O (that the animals are zebras, that you have hands). In particular, we give the same kind of negative explanation for the two cases: We seek to explain why it can seem that you don't know that $\sim H$ in both cases by claiming that the discussion of whether one knows $\sim H$ has a tendency to put into place high standards that we don't meet with respect to our belief that $\sim H$ (nor with respect to our belief that O).

But when it comes to our positive explanations, the matter of *how we might have come to have* this knowledge-by-relatively-low-standards of $\sim H$ looks quite different

¹⁴ See Holiday's use of "epistemic work," and also of "empirical work" and "epistemic achievement" (Holliday 2015: 113–17).

for the different kinds of skeptical hypotheses.¹⁵ For, though a lot of it is quite old and operating in the background of our thinking, we do seem to have empirical evidence in virtue of which we are at least quite reasonable in believing that the animals in the zoo are not cleverly painted mules. It may be hard for us to articulate just what the evidence we have here is or how it supports our belief, but we've been around the block a few times, and that the world at least tends not to work in such a way that those animals would be cleverly painted mules seems like something we have learned, in what is probably a very complicated way (this is probably in some ways an impressive, even if utterly common, "epistemic achievement") from our experience of the world. These complicated, backgrounded reasons we have for thinking $\sim H$ in this moderate case would *appear* to be insufficient for *knowledge* of $\sim H$ (even if $\sim H$ is in fact true), but addressing this appearance of ignorance is the negative explanatory task that has been the focus of us contextualist Mooreans. In the case of such moderate skeptical hypotheses, we hold that the evidence in question, which *appears* to be only good enough to enable us to have reasonable belief, and not knowledge, in fact allows us to know that $\sim H$, at least by uninflated standards for knowledge (and of course, when we are in fact right that $\sim H$), and we employ our negative explanation to account for why there is an appearance of ignorance here.

Things are quite different with our knowledge_o that we're not BIVs (and that other "global" or "radical" skeptical hypotheses are false). That at least does not seem to be

¹⁵ Here I follow the insightful lead of Stewart Cohen (1988: 111) in drawing this key distinction and giving quite different treatments of our knowledge_o of the falsehood of moderate vs. radical skeptical hypotheses. However, Cohen might not agree with my following discussion of how I think we know_o that moderate skeptical hypotheses are false. Cohen wrote of us knowing by means of "statistical" grounds. I don't think that that's quite right, so far as I understand the diagnosis. As you are about to see, I think this knowledge_o tends to be based on old and backgrounded empirical evidence, rather than, at least generally, evidence that is in a good, enlightening sense, "statistical" (though it is evidence which does not entail the conclusion it is supporting, so if non-entailing is all one means by "statistical," then I agree that our knowledge_o that moderate skeptical hypotheses are false is in that sense based on "statistical" grounds). Still, I am following Cohen's general strategy here, the main idea of which is to explain why the empirical evidence we have for the falsehood of moderate skeptical hypotheses appears to be insufficient for knowledge.

Dogramaci (2014) argues that "rationalism" (which in his use requires antecedent justification for $\sim H$ if one is to know that O) cannot handle certain "one-off" possibilities of error that he devises, on which one is generally reliable in perception, but for some reason is deceived about O on the occasion in question. Some of Dogramaci's scenarios would seem to be of the moderate kind that I hold we know_o not to obtain on the basis of empirical evidence that does not *appear* sufficient to underwrite knowledge of $\sim H$, but that in fact, according to me, is sufficient for at least knowledge_o of $\sim H$. Crucial to Dogramaci's case is his claim that one cannot base knowledge on "statistical evidence," and his closely related claim that we don't know that we lose the lottery (in standard lottery cases), so my argument in Chapter 5 that the best solution to the puzzle of conflicting intuitions concerning knowledge in lottery cases is one on which we do know_o that we've lost the lottery will be key to the differences between us here. However, other of Dogramaci's skeptical hypotheses (like that I'm a one-off victim of an evil genius) would be of the radical kind that we don't have effective empirical evidence against, and that I think we know_o a priori to be false. About those cases, the key is that I do not, as Dogramaci's targets seem to, construe this knowledge as being derived from premises about my reliability. Rather, as we'll see, I think we come to know_o that we're not in radical skeptical scenarios because those hypotheses conflict with our perceptual beliefs and the picture of the physical world we build on their basis.

the kind of thing that we could have in any good sense learned, or even come to reasonably believe, based on evidence we have gathered from our experience of the world, as all the evidence otherwise suited to that purpose would seem to be *undermined* (as it's extremely plausible to say) for the purpose of establishing or even supporting conclusions of the likes of that we're not BIVs. So, to cut right to the chase here, it seems that any knowledge_o we might have of such things would have to be in some sense a priori. The position that our knowledge that radical skeptical hypotheses are false is a priori has been called "rationalism" (Dogramaci 2014; Willenken 2015), but we will call it "a prioritism," in order to be able to use "rationalism" to designate an old, wide-ranging philosophical tradition that isn't tied down to a position on the particular issue of how we know that radical skeptical hypotheses are false.¹⁶

So, we contextualist Mooreans, and most clearly Stewart Cohen and myself, have adopted a contextualist version of a prioritism, holding that our knowledge_o that we're not BIVs is a priori.¹⁷ And this a prioritism has seemed to many to be a major problem with our view. Indeed, this seems problematic not just to critics, but also to Cohen himself, who describes this aspect of our position as "distasteful," and as something he is "not entirely happy with," but also, in our skeptical predicament, as the "bullet [he is] most prepared to bite" (Cohen 2000: 105–6).

¹⁶ Non-contextualist a prioritists/rationalists would include Crispin Wright (esp. 2004) and Roger White (2006: esp. 552–3), though see Holliday's (2015: 114, n. 34) case that Wright and White don't go so far as to claim that what we possess a priori is actually knowledge (and so, I suppose, that they should just be counted as accepting a prioritism lite).

¹⁷ Cohen (2000) discusses this aspect of his position explicitly, and as we're about to see, he seems rather dismayed by it. But this element of his thinking goes way back: It would seem to be what is behind his claim that it is "intrinsically rational" for us to believe that radical skeptical hypotheses are false in Cohen (1988: 112). Though perhaps a bit half-heartedly, I leaned toward a prioritism in DeRose (2000a: 138): "I suspect that the best ways of filling out and then evaluating my alternative, contextualist Moorean account of how we know that we're not BIVs will have it come out also as an account according to which our knowledge that we're not BIVs is a priori." I was cautious; labeling this only as a "suspicion," mainly due to questions about just how to understand "a priori" that affect how the term applies to my views. We will be able to explore this suspicion in what follows in this chapter. Though to the best of my knowledge they never explicitly claimed that our knowledge_o that skeptical hypotheses are false is "a priori," what Stine and Lewis wrote about how we know that irrelevant alternatives are false can seem (and does seem to me) to point (perhaps vaguely) in that direction. I have in mind especially what they wrote at and around the following passages. Stine: "I do know that it is not a mule painted to look like a zebra. I do not need evidence for such a proposition. The evidence picture of knowledge has been carried too far. . . . [I]f the negation of a proposition is not a relevant alternative, then I know it—obviously, without needing to provide evidence—and so obviously that it is odd, misleading even, to give utterance to my knowledge. And it is a virtue of the relevant alternative view that it helps explain why it is odd" (1976: 258). Lewis: "Do I claim you can know P just by presupposing it?! Do I claim you can know that a possibility W does not obtain just by ignoring it? Is that not what my analysis implies, provided that the presupposing and the ignoring are proper? Well, yes" (1996: 561–2). As I mentioned in note 15, following Cohen (1988: 111), I think we must distinguish radical from moderate skeptical hypotheses, and give different treatments to each on just this point, and Stine and Lewis did not do this. Still, it's in virtue of features of our views that we all share as contextualist Mooreans that Stine and Lewis are at least pushed in the direction of a prioritism.

So in spelling out how we know_o that we're not BIVs, I will be paying close attention to whether and in what ways this knowledge_o is a priori on my view, and whether this is a problem for my account. It turns out that my attitude is very different from Cohen's. Sometimes in philosophy you have to bite bullets and accept distasteful positions in order to avoid what seem to you even worse problems, and it is good to be upfront about it, as Cohen is, when that is what you are doing. But in the matter of the a priority of our knowledge_o that radical skeptical hypotheses are false, the problem seems to me to dissipate upon closer inspection. Though, as we'll see, it's no easy call for me whether to rule our knowledge_o that we're not BIVs a priori, that turns out to be just a tricky matter of how to use the murky, highly theoretical, term, "a priori," and I never see myself as being boxed into any substantive position where I have to be doing anything that might be likened to biting bullets.

One tool for keeping an eye on this issue is addressing the following variant of our basic skeptical argument, AI, which simply breaks AI's first premise into two cases (where we resolve that, however exactly we understand "a priori" and "through experience," we will coordinate our construals of those two terms so that it will be true that all knowledge one has must be either one or the other, thereby avoiding the need for an explicit premise to that effect):

Split Cases AI (AI-SC)

- 1a. I don't know a priori that not-H.
 - 1e. I don't know through experience that not-H.
 2. If I don't know that not-H, then I don't know that O.
- So, C. I don't know that O.¹⁸

I present AI-SC without a subscripted o attached to each occurrence of "know," because we will want to be thinking of what to say about this kind of skeptical argument in general. But the interesting dilemma it presents for me and other contextualist Mooreans (that goes beyond the problems inherent in regular, old AI) is what to say about this new form of the argument when the occurrences of "know" are all taken to express knowledge by ordinary standards (knowledge_o).

It turns out that my views in the epistemology of perception changed some years ago in ways that would render our bits of knowledge_o that radical skeptical hypotheses are false somewhat worse candidates for being counted as a priori, and so (further) cloud the matter of which of 1a or 1e I reject, where the occurrences of "know" are understood as being set to ordinary standards. However, I will engage in a bit of pretense, and, except in a note,¹⁹ I will write as if I held my old

¹⁸ Thanks to Brian Weatherson for suggesting that I consider such a skeptical argument. For some purposes, this argument can be seen as intermediate between AI and the skeptical argument that Pryor presents and that we discuss in Appendix A.

¹⁹ In DeRose (2005), I defend the new (for me—similar positions have been held by others) view, which I call "Direct Warrant Realism" (DWR), as opposed to the simpler view I am here pretending to still hold,

view, on which it's a bit clearer that our knowledge_o of ~BIV is best construed as a priori and it is 1a that should be rejected. The reason for this procedure is that even my current view would assign an a priori aspect to our knowledge_o of ~BIV that would disturb those who object to a priori knowledge of such things, and so the change really does not seem to affect the important issues in front of us, but would only complicate my handling of it. My old view makes for a cleaner contrast with the views I oppose. I think that an unfounded fear of ruling certain types of knowledge to be a priori has driven a good deal of thought about skepticism off-track. The notion of a priori knowledge seems to me to be quite flexible, and on various views, including both my old and my current view, it isn't completely obvious whether to understand the notion in such a way that our knowledge_o of ~BIV should count as a priori—though on what I'm inclined to think is the best way to draw the distinction, things are as I intimate above: It's a priori on my old view, but is at least not so clearly so on the new. Still, what's important in responding to AI-SC is to see how tricky some of these calls are, and to avoid quickly grasping on to simple formulas that would ban certain types of a priori knowledge. It proves best for making the needed points to adopt my older, simpler, and bolder view.

which I call "Stringent Direct Realism" (SDR). On DWR, simple perceptual beliefs enjoy direct warrant (warrant that does not come from being based on other beliefs), but not enough for rational acceptability or knowledge. They are then rendered justified/rationally acceptable by the added indirect warrant they derive from the mutual support they provide to one another in virtue of how they fit together to form a coherent view of the physical world. I admit/accept that our perceptual beliefs are typically formed in what can well be called a *psychologically* "immediate" way: They are formed without giving any conscious thought to the relations, of coherence or otherwise, the belief we are forming bears to other of our beliefs. But, I continue: "I *still* find DWR to be a very attractive account of the warrant of perceptual beliefs. This is largely because I think It's wrong to require that a belief be formed due to a conscious noticing of its relations of support to other beliefs in order for those other beliefs to transmit warrant to it. Perhaps, though we take no notice of a perceptual belief's relations of 'coherence' (relations of mutual support) to other perceptual beliefs as the belief in question is formed and maintained, we are appropriately *sensitive* to its coherence with these other beliefs, where such sensitivity consists in such facts as that we wouldn't hold or continue to hold the belief in question, or at least wouldn't continue to hold it to the degree that we do, were it not for its coherence with our other beliefs. One could be in that way sensitive to the coherence of one's beliefs even where one gives no conscious thought to the relations in question, and yet it seems to me that warrant might very well be transmitted among beliefs in virtue of the believer's sensitivity to the evidential relations that hold among them, even where she gives no conscious thought to those relations" (DeRose 2005: 160). See DeRose (2005) for a defense of DWR from a variety of objections. On DWR, it is crucial to the justification of one's perceptual beliefs that one's experience has turned out to follow a path on which those perceptual beliefs display the needed positive coherence with one another. This (to now address our current concern) could introduce an empirical element into our justification for believing that radical skeptical hypotheses are false that could, depending on how we use the term, ruin the claim of our beliefs in the various ~Hs to being items of "a priori" knowledge_o. However, even when accepting DWR, I still assign an important a priori element to our justification for believing that radical skeptical hypotheses are false that will both disturb opponents of a priori knowledge of (deeply) contingent truths but also encourage us to construe our justification for these ~Hs as not being dependent on seeming problematic inferences from beliefs in the likes of O, nor from fancy philosophical arguments for the falsehood of the Hs.

7.11. The Basic Easy Account of How We Come to Know_o that We're not BIVs

Since we are seeking a non-heroic, easy account on which the non-philosophical, without the aid of any fancy philosophical arguments, know_o that they are not BIVs (or at least could or would come to know_o that ~BIV were they to consider the question), a good place to start our thinking is with this question: Whether or not the resulting belief constitutes knowledge, how do we (or at least those of us who ever consider the matter) generally come to think that we're not BIVs? I take it that we do come to think that before, and independently of, any fancy philosophical arguments we devise or learn of to the conclusion that we're not BIVs. And I also take it that the answer to our question, in short, is that when we consider the BIV hypothesis, we find it bizarre, reject it (at least as what we should or will believe here), and in fact find ourselves believing that it is false.

It may be worth saying a bit more about the phenomenology of this process, though I don't think the fine details of this phenomenology are important to how we should end up evaluating the beliefs that result. But it is often held that there is a characteristic phenomenology to the way we come to have a priori knowledge of necessary truths, like, for instance, bits of arithmetic: See, for instance, Alvin Plantinga's description of this phenomenology (Plantinga 1993: 104–6). And it is widely held that this phenomenology involves coming to think or its seeming to us that the truth being apprehended is *necessarily* true (Plantinga 1993: 105). I will be claiming that the knowledge we gain of ~BIV by the process I briefly described in the previous paragraph should also be construed, at least on the best sense I can make of the notion, as a priori. Yet the process by which we come to believe ~BIV does not involve coming to think that the proposition believed is necessarily true.

However, the process I am characterizing does seem to have its own distinct phenomenology, described by Thomas Reid, who made the following observation to aid in discerning “first principles” (things reasonably believed, but not on the basis of argument), which Reid thought often were, and were felt to be, only contingently true:

We may observe, that opinions which contradict first principles are distinguished from other errors by this; that they are not only false, but absurd: And, to discountenance absurdity, Nature hath given us a particular emotion, to wit, that of ridicule, which seems intended for this very purpose of putting out of countenance what is absurd, either in opinion or practice. This weapon, when properly applied, cuts with as keen an edge as argument.²⁰

I don't think Reid's above description applies to all of our immediate knowledge, but, following Reid, we may observe that there seems to be a particular way that certain suggestions, including ones that are only contingently false, like that we are BIVs,

²⁰ Reid (1785: 462). Michael Bergmann in his work has drawn attention to this nice passage from Reid (for instance, at 2004: 723).

strike as absurd, ridiculous, bizarre, or far-fetched. And my suggestion is that we come to believe that \sim BIV through the proposal that we are BIVs so striking us.

Though I think we do best in identifying what is distinctive in the phenomenology involved by focusing on the negative side of things (i.e., on how what we come to reject seems to us), as I have just done, looking briefly at the positive side, and following Geoff Pynn, we might say that there is a particular way that what we come to *accept* through this process (\sim BIV, in the case we are particularly interested in) seems *obvious* to us (Pynn ms.a). To the suggestion of BIV, we are moved to exclaim: "Absurd!" "Outlandish!" "Aw, come on!" "Get out of here!" "Come off it!" "That's ridiculous!" To \sim BIV: "Of course!"

On our easy account of our knowledge_o of \sim BIV, someone who in normal circumstances comes to believe \sim BIV in the common, simple way we have been considering, without any support from any fancy, philosophical argument, thereby comes to know_o that she is not a BIV. (And similar stories account for how they would come to believe that other radical skeptical hypotheses are false.)

On our double-safety account of knowledge, as it has so far been articulated, a belief arrived at in that way will amount to knowledge if (it is true and if) there are no too-nearby worlds in which the subject's belief on the matter of whether she is a BIV is false. We might pause briefly to note how our account of knowledge connects with some of the phenomenology of how we come to hold the relevant beliefs, since some of the ways we are led to belittle the skeptical hypotheses employ related spatial metaphors, as when we say that they are too "far-fetched," "far-out," or "remote" to be worth worrying about. And, based on how we think of the kind of "distance" that's relevant here, and our boringly normal view of the situation most people in fact are in, the non-skeptical safety theorist will hold that their safety condition, at least at standards _o, is typically met by our belief that \sim BIV, when it is arrived at in the normal way.

7.12. Epistemic Justification, Epistemic Conservatism, and the Basic Easy Account of How We Come to Know_o that We're not BIVs

But how is such a belief *epistemically justified*—justified in a way required for knowledge?²¹ Some safety theorists may think that no type of justification is required for knowledge, or that it turns out that their safety condition, externalist as it looks,

²¹ I use "epistemically justified" in a way that it picks out a notion of justification that is internalist and very relevant indeed to the notion of knowledge. My basic attitude toward "justified" has always been in line with these observations by Steven L. Reynolds:

"Justified," as applied to belief, seems a bit like philosophical jargon, although it is not usually regarded as a technical term. Our intuitions about it are less definite, and so perhaps more open to the influence of philosophical theory, than the intuitions we have about knowledge. I propose to use it as a technical term, explicitly defining it in terms of knowledge. That will allow us to set aside distracting subtleties about its ordinary uses, to

nonetheless somehow manages to articulate the (only) kind of justification that is needed for knowledge. But there is a suitably internalist account of our justification for \sim BIV that fits in very well with the easy, basic safety account of our knowledge of \sim BIV: One could appeal to a suitable form of *epistemic conservatism*, according to which, absent good reason to think things are otherwise, one is justified in believing what seems to them to be the case.²² (As conservative theories are refined, they may start limiting the seemings to which their theory can be applied; a “suitable” form of conservatism here is one that would allow application to \sim BIV, where that belief is arrived at in the typical, simple way.) We don't need any heroic proofs that we're not BIVs: Absent good reasons to think we actually are BIVs,²³ we are justified in thinking that we are not. Indeed, I was inclined to so appeal to epistemic conservatism to combat skepticism about epistemic justification before I came to accept any kind of safety account of knowledge.²⁴

focus instead on our use of the term “know,” which I take to be the central epistemic evaluation for the Cartesian skeptical arguments. (Reynolds 1998: 533)

I use “epistemically justified” to mark what I call my “semi-technical” use of “justified,” which is much like Reynolds's proposed use of “justified” (which use was, before the label was popular, a fine example of the kind of “knowledge first” epistemology championed by Williamson and his followers), so far as the above quotation goes. Reynolds continues:

Justification, let us say, is how knowledge appears to the knower, how it seems “from the inside.” More formally: S is justified in believing that p if and only if it really appears to S that he has knowledge that p. (1998: 534)

In my use of “epistemically justified,” I follow Reynolds in the basic idea that it is what knowledge is like from the inside. But I don't follow Reynolds in his use of the notion of the appearance of knowledge to implement the basic idea. (But those interested in that attempted implementation should see the later Reynolds 2013.) Rather, I like to use the notion of “internal twins” (for whom everything seems the same “from the inside”) to locate my knowledge-based use of “epistemically justified,” and say that, at least for most ordinary, contingent propositions p, if S has the attitude of knowledge toward p, then S is epistemically justified in having that attitude toward p if and only if some possible internal twin of S knows that p. (Ichikawa (2014) is an account of epistemic justification in the vicinity of this. The limitation to “most ordinary, contingent propositions” is due to the obvious problem of applying such an account to justified beliefs in necessary falsehoods, and related problems, pointed out to me by Jeremy Goodman, with some contingent propositions. Complications will arise in cases where the “external” differences in the twins' settings cause their corresponding beliefs to differ in content.) I am here depending on there being some good use of “justified” that, at least roughly, answers to this characterization. Those who are suspicious of “internalist” notions of justification should be suspicious of my use of “epistemically justified.” My own suspicions have always run the other way: I've always found it very difficult to find any good use of “justified” such that, say, one of my beliefs and the corresponding belief of my BIV-internal-twin could differ with respect to their justificatory status (with respect to whether and to what extent we are justified). I always suspect that while such “externalist” differences can be very important to the epistemic evaluations of the beliefs in question (they can of course make the difference between one twin knowing and the other not), they can't mark out any good sense in which one twin's belief is “justified” while the other's is not. Something like Conservatism can then be put forward, as I am doing here, as a thesis about epistemic justification, so characterized.

²² Perhaps the most important recent defender of epistemic conservatism is Michael Huemer (see esp. Huemer 2007, and then his later developments on his view).

²³ Later in this chapter, in Section 7.20, we will consider a case where one's justification for \sim BIV gets defeated by one coming to have such reasons to think one is a BIV.

²⁴ My conservatism surfaced first in my exposition of the great epistemic conservative, Thomas Reid (DeRose 1989: 326–31), and was subsequently modified a bit and given in my own voice in DeRose (2005).

Drawing the various elements together, our positive account of our knowledge_o of ~BIV now looks like this: When we consider the BIV hypothesis, we find it bizarre and reject it (at least as what we should or will believe here) and in fact thereby come to believe that the hypothesis is false. That belief of ours is epistemically justified in virtue of its seeming to us to be true, if we have no good reason to think that we are BIVs. And it then amounts to knowledge_o if it is true (if in fact we are not BIVs) and if there are no worlds that are too nearby by standards O in which we are wrong about whether we are BIVs. And since we have no good reason to think we are BIVs, and we in fact are not BIVs, and there are in fact no worlds that are too nearby by standards O in which we're wrong about whether we're BIVs, we do know_o that we're not BIVs.

Would such knowledge_o of ~BIV be a priori? As our account is *so far* formulated, it looks like it would be, for it does not look like there is any point in our account at which justification derived from perceptual experience is playing a role in our coming to justifiably believe and know that ~BIV. But there is potential for things to appear differently upon a closer look at the normal process by which we come to think that we're not BIVs.

7.13. A Prioritism vs. Dogmatism

In particular, we can dig a little deeper into the normal process by which we come to believe that we're not BIVs by asking *why* we find the BIV hypothesis so rejectably bizarre. Perhaps it strikes us in such a way because it conflicts sharply with our standing, well-justified perceptual beliefs (and the highly justified view of what the world is like that is built up from those perceptual beliefs). In that case, our justification for ~BIV may be derived, directly and/or indirectly, from our justification for the likes of O, with the latter serving as our evidence for the former. And since our justification for the likes of O is presumably derived from experience, our justification for ~BIV, which on this picture we derive from our justification for the likes of O, will be empirical as well, and so won't be a priori, after all. In that case, perhaps epistemic conservatism should be applied directly only to our simple perceptual beliefs, as an account of how they come to be justified in the process of perception (and so presumably would not be a priori), which beliefs then form our (in that case, a posteriori) basis or evidence for believing that we're not BIVs.

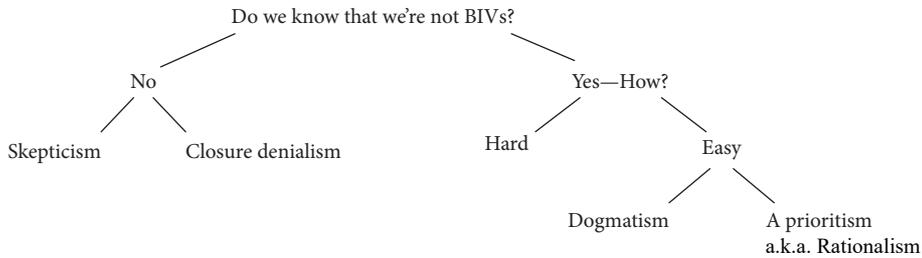
This would bring us to an account of our justification for ~BIV that I will call "dogmatism," not as a term of abuse, but because it is the position at least suggested in James Pryor's "The Sceptic and the Dogmatist" (Pryor 2000). Pryor labels his own position "dogmatism," and I suspect that as *he* uses the term, I count as a "dogmatist" as well, but I will be using the term more narrowly, so that one counts as a "dogmatist" only if one takes the position that we do know ~BIV (or at least can know it), and that we base such knowledge on our perceptual beliefs about the external world. Though I don't find Pryor to be entirely clear on the matter, I find

such (narrow) dogmatism to be suggested primarily by his apparent agreement with this premise of the skeptical argument he formulates:

- (5) Either you don't know you're not being deceived by an evil demon; or, if you do know you're not being deceived, it's because that knowledge rests in part on things you know by perception. (Pryor 2000: 524)

Pryor's apparent acceptance of (5), together with his warmth toward at least moderate forms of closure and his anti-skepticism, would suggest that he thinks we know (and justifiably believe) that radical skeptical hypotheses are false, or at least can know (and justifiably believe) these things if we consider the matter, and that such knowledge (and justified belief) "rests" on the likes of O. At any rate, I find such a position more interesting than a closure-denying one, so that's the position I will be calling "dogmatism."

Though such "dogmatism" is a fairly natural view (in a way I tried to draw out in the first paragraph of this section), I deny it, instead holding, as I've indicated, to a prioritism. So our options look like this:²⁵



I find myself in general agreement with a lot of what seems to be driving Pryor's views in the epistemology of perception. (As I've already indicated, I suspect I, like Pryor, am a "dogmatist" as he uses the term.) What accounts for our difference over how we know ~BIV—or at least my difference with the position I take to be suggested

²⁵ In light of considerations raised in note 19, we may need to add a third "easy" option, which we might call "weak a prioritism" (and we might then change the name of "a prioritism" to "strong a prioritism"), for those who, as I'm inclined to, accept a DWR rather than an SDR account of the justification of our perceptual beliefs.

John Greco proposes another easy option (Greco 2002: 560–3): That we know we're not brains in vats because we perceive that that is so—and, to get specific about sense modality, we see that it is so. That *sounds* unpromising to me. As Pynn nicely puts it: "I think this is no more plausible than saying that I can see that a bill isn't a perfect counterfeit" (Pynn ms.a.). As Pynn's quip insinuates, there's a (roughly) insensitivity problem with Greco's claim: It seems wrong to claim to see that p is the case (and to thereby come to know that it is so) when things look just as they would if p had been false. (I also, and only semi-jokingly, worry about the choice of sense modality here, and wonder whether it means that those who have always been totally blind don't know that they're not BIVs.) Such an intuitive problem could perhaps be explained away. Much of our discussion to come (starting in Section 7.15) of whether we know ~BIV through experience or rather know it a priori, would be relevant to evaluating Greco's proposal, since coming to know ~BIV by seeing/perceiving that ~BIV would be a way of coming to know it through experience. My contention that ~BIV is something we bring to, rather than derive from, experience would push against Greco's proposal.

in Pryor's work? I think it's fair to say that we are driven apart by our different aversions. We seem to share aversions to bold skepticism and to closure denialism, but Pryor's aversion that I do not (at all) share is to ruling that knowledge of the likes of that we're not BIVs is a priori. It is an aversion that runs deep in traditional western epistemology, and that I will be wrestling against starting in Section 7.15.

7.14. Problems for Dogmatism

My aversion is to thinking that the likes of O can serve as evidence by which to know that radical skeptical hypotheses are false. For while our perceptual beliefs seem to be fine bases for various inferences we might make, as I have already noted, they seem to be "undermined" (in a very intuitive use of this term) for the purpose of establishing something like ~BIV.

To revisit a hopefully useful analogy I've used elsewhere (DeRose 2000b: 704–5; see the similar case at Vogel 2000: 613–14), suppose I hear on the radio that "The Cubs beat the Cardinals," in a list of results that were read during the sports segment of the news, and that that radio report is my only source of information for that result. Here, at least under suitably favorable circumstances, that the Cubs beat the Cardinals is evidence I have, or at least seem to have, and can properly cite, in support of some conclusions. If, for instance, you claimed that the Cubs haven't beaten any good teams this year, I seem to be in a position to argue, "Well, they beat the Cardinals just today, and the Cardinals are a good team." However, that the Cubs beat the Cardinals seems not to be properly usable as evidence for other conclusions (to which it would be very relevant if only I were in a position to cite it). Suppose that for some reason—or perhaps for no good reason—you get it into your head to discuss the issue of whether the sports segment I heard on the radio was (at all) accurate or reliable in the scores it reported. That the Cubs beat the Cardinals seems not to be evidence I have, or at least seems not to be evidence I can properly use, for the purposes of arguing for a pro-reliability conclusion here, given how I came to hold that evidence. If I were to argue, "Well, it certainly seems that the report was reliable, because, according to it, the Cubs beat the Cardinals, and they did, and it reported that the Marlins beat the Mets, and they did, and . . .," you shouldn't be very—or, really, at all—impressed with this "track record" argument if you knew that in each case, my only source for the "and they did" part was the very radio report whose accuracy was presently in question! That the Cubs beat the Cardinals, and the like, seems not to be evidence I have, or at least not to be evidence I'm allowed to use, for such a conclusion as that the radio report was reliable, at least in these circumstances (I have no other sources of information for the results of yesterday's games). This evidence, which I seem to have and am allowed to use when it's another issue that's in question, seems to be undermined for the purpose of establishing that the radio report was accurate. And this doesn't seem to be just a point about proper dialectical practice when arguing with an opponent: It seems equally illegitimate to base a belief

in the reliability of the radio report on such evidence even when this “track record” argument occurs privately in my own thought.²⁶ I do need to be justified in believing the radio report is reliable to be justified in believing the content of what it reports, or at least so it seems to me. But it also appears that this justified belief in the reliability of the report can't be evidentially based on my beliefs in the scores themselves, when my only source of information for those scores is the report in question. Presumably, my justification for believing that the report is reliable derives, in a complicated way, from my experience of the world, from which I have learned that the world tends not to operate so that such reports, in such circumstances, would be inaccurate. (The belief that the radio report is reliable will be an insensitive belief under the

²⁶ This track record case is an example of what, following William Alston, has come to be called an “epistemically circular” argument (Alston 1986). At DeRose (1992b: 230–1), I defended epistemically circular arguments, at least to the point of thinking it is not implausibly uncharitable to ascribe such reasoning to Descartes in the *Meditations*. And I suppose I still think that. However, I have (further) soured on epistemic circularity since writing that paper. This (still) seems right: “Next, as Alston has pointed out (Alston 1985, p. 449), we should remember that the use of a faculty could result in one's coming to learn that the faculty is unreliable. Alternatively, it might issue no result regarding its own reliability. In light of this, if Descartes were right that his faculty of clear and distinct perception is self-verifying (rather than self-undermining), this would by no means be an obviously worthless result” (DeRose 1992b: 230–1). But I now realize that we must distinguish the epistemically circular argument itself being a good argument, specifying a line of reasoning that could render it more reasonable to believe its conclusion, from there being persuasive force to a *related but distinct* argument that the epistemically circular argument suggests and that utilizes a premise to the effect that self-verifying faculties have some sort of virtue, together with a premise that the faculty being discussed is self-verifying. Once this distinction is made, I find this defense of epistemically circular arguments to lose its force.

Michael Bergmann has prominently defended epistemically circular reasoning (Bergmann 2004), arguing that anyone who accepts the foundationalist thesis that “there can be noninferentially justified beliefs” should accept the legitimacy of epistemically circular track record arguments of just the kind I'm rejecting (2004: 713–17). In brief, I think the way to resist Bergmann's argument, while retaining the foundationalist thesis and eschewing skepticism, is to hold that where one does justifiably believe the premises of these track record arguments, as one often does, one must and does already have justification for the conclusion. One does not base one's beliefs in the premises on an epistemically prior belief in the conclusion, so the premises can be noninferentially justified—as they are in the cases Bergmann is interested in. But neither can one's justification for the conclusion be based on the premises in these cases. The justification for the premises and for the conclusion are epistemically “simultaneous,” as it were, neither prior to the other. Justification then “fails to transmit” from the premises to the conclusion because that justification is, and must be (if one is justified in believing the premise), already there (at the conclusion). The specific place where I think Bergmann's argument goes wrong is at his transmission principle (c) (2004: 715), which I think needs some additional “unless” clause, exempting from its scope cases where premises are “undermined” for the purpose of defending the conclusions they are being used in support of, or alternatively (and probably preferably) exempting cases where one already has justification for the conclusion. (Bergmann does specify in his principle that the conclusion is “believed for the first time,” but one can have justification for a claim one does not yet believe.) Bergmann, who is more friendly than I am to reasoning that seems epistemically circular, will likely not see the need for any such an added clause. But I think just about anybody who is not already friendly to epistemically circular reasoning will likely insist on some such weakening of the transmission principle, so an argument that rests with all its weight on the unweakened principle is not likely to win over anyone who is not already on board with epistemic circularity. I think we should be very strongly motivated to resist Bergmann's argument, since the epistemically circular track record arguments he is defending seem to me transparently bad—and not just in Bergmann's “questioned source contexts.”

circumstances in question, so recall that I will expect it to not seem to amount to knowledge, and have an explanation for this appearance.)

In what seems a similar way,²⁷ that I have hands *appears* to be *undermined* as a piece of evidence if the issue under discussion is whether or not I'm a BIV: It *appears* not to be legitimately usable as a piece of evidence for such a conclusion.²⁸ And likewise then for my other perceptual beliefs that put my picture of the world at odds with the BIV hypothesis.

How shall we proceed if the appearance of trouble is accurate, as I suppose, and there is a genuine problem with using the likes of O as evidence for \sim BIV? While I think this would make our situation with respect to these beliefs like the scenario we considered involving the Cubs' victory and the reliability of the radio report in some important ways, there is also an important difference between the cases. For my belief that I'm not a BIV (as opposed to my belief in the reliability of the radio report) doesn't seem to be the kind of thing I might then have *any* grounds for believing that are derived from my experience of the world—because *all* the potential bases for this belief seem undermined. And that, again, is precisely our problem here: If I can't come to know_o that \sim BIV through the likes of O, then it seems I can't have learned it through my experience of the world *at all*. But then, skeptics, closure denialists, and dogmatists alike will ask, perhaps quite pointedly, how *can* I have come to justifiably believe and to know_o that I'm not a BIV?

And, again, the answer would seem to (have to) be: In some good sense, I know_o it a priori—which would explain why someone like Pryor, with a strong aversion to thinking there could be a priori knowledge of such things, would end up instead (to use the words of Cohen, who long ago perceptively discerned the nature of the choice before us here, but adding the “dogmatist” label that Pryor has since made popular) biting the other bullet and “endorsing what looks to be objectionable [dogmatist] reasoning.”

²⁷ I wouldn't count *I have hands; so, I am not a BIV* as an epistemically circular argument, as I'm inclined to use “epistemically circular” narrowly, covering only arguments whose conclusions are explicitly to the effect that some way of forming beliefs is reliable, and (some of) whose (needed) premises are derived from the way of forming beliefs in question. (And I don't use the term so that it's analytic that “epistemically circular” arguments are defective: The terminology, as I use it, leaves it open whether there's anything wrong with “epistemically circular” arguments.) But my claim here is that the proposed anti-BIV argument suffers from the same kind of problem that plagues the epistemically circular track record argument in our baseball results case.

²⁸ I put this forward as a mere appearance because, while I hope my discussion to this point has managed to tap into the “undermining” intuitions involved here, I don't take myself to have properly argued for the conclusion that our perceptual beliefs can't be properly used as evidence against the BIV hypothesis. We are here dealing with problems for Pryor's dogmatism of the same genus as some of those that are pressed by Roger White in his paper (to which I'm tipping my cap in the title I chose for this section) “Problems for Dogmatism” (White 2006: esp. 543–52)—though see also Pynn (ms.b), who presses such problems in a way especially congenial to my views, and Cohen (2000: 105–6), who saw long ago that “Our options seem to be accepting contingent a priori knowledge or endorsing what looks to be objectionable reasoning” (2000: 106), but also Pryor's response to White (Pryor 2013). I take these problems for dogmatism to be real and quite serious. My goal is not to solve them, for I take them to be insoluble; nor to further advance them; but to get around them, by taking a non-dogmatist (in our narrow sense) position.

7.15. Deeply Contingent A Priori Knowledge: An Intolerable Problem for A Prioritism?

So, what is so bad about supposing our knowledge_o that we're not BIVs is a priori? The aversion seems to be rooted in the contingency, or the "deep contingency," of the fact that we're not BIVs. The aversion to a priori knowledge of contingent truths goes way back in the history of philosophy; the "deeply" qualification represents a relatively recent (late twentieth century) sharpening of the aversion, made in response to examples of supposedly a priori knowledge of contingent facts given by Saul Kripke. The terminology used in the qualification is due to Gareth Evans, who responded to Kripke's examples by distinguishing between two classes of contingent statements, the "deeply contingent" and the "superficially contingent" (Evans 1979: 161); allowing that Kripke had shown that there can be a priori knowledge of merely superficially contingent statements; but denying that there is knowledge of deeply contingent statements (Evans 1979). We won't need to worry about whether Kripke really provided examples of a priori knowledge of contingent facts, nor how to draw this distinction (nor the related matter of the slippage between my writing of a priori knowledge of facts or propositions while Evans puts matters in terms of the knowledge of statements), as we won't be dealing with any cases of merely superficial contingency: I will be defending the a priori nature of our knowledge_o of the falsehood of radical skeptical hypotheses despite the deep contingency of the non-obtaining of those hypotheses. (Indeed, in what follows, though I will sometimes add the "deep[ly]" qualification as a reminder, wherever I use "contingent," even without a reminder, I should be understood as referring to the *deeply* contingent.)

Though Evans and those that follow him may have narrowed the scope of the traditional aversion, its force seems undiminished, Evans himself forcefully and memorably declaring that "it would be intolerable for there to be a statement which is both knowable *a priori* and deeply contingent" (1979: 161).

But why think that? Setting aside accounts of the a priori that would take it as just part of what "a priori" means that only necessary truths can be known a priori, and accounts on which, simply as a matter of the meaning of "a priori," a belief can't count as a priori unless the process by which it is formed produced in the believer the appearance of necessity, it turns out that expressions, including extremely forceful ones, of the aversion are much easier to find than are explanations for *why* the forbidden thought is so intolerable.

In a paper in which he goes on to seriously question Evans's exclamation, John Hawthorne writes that "It is clear enough what prompted this remark by Evans." Hawthorne explains:

Suppose that having understood some sentence *s*, one does not thereby obtain some guarantee of a verifying state of affairs. One will in that case find it perfectly conceivable that the actual world enjoys a distribution of objects and properties that falsifies *s*. But now, it seems, one will

need to do some empirical investigation to figure out whether the actual world is a verifier or a falsifier of *s*. There thus appears to be a straightforward argument against the possibility of deeply contingent a priori knowledge. (Hawthorne 2002a: 248)

But if that were all that could be said in defense of the prohibition, it would be hard to see why it should be accepted at all, much less with the gusto with which Evans endorses it. For as Hawthorne quickly points out, the argument he has ascribed to those who would oppose a priori knowledge of the deeply contingent “seems to rely on the principle that if I can conceive that not *P* is actually the case, then I do not know that *P*.” And this is an extremely dubious principle: “But it is generally agreed that a knowledge-conferring warrant for some *P* need not offer a watertight guarantee of *P* and thus need not render it inconceivable that the actual world falsifies *P*. Why then require such a guarantee when it comes to [a] priori warrant?” (2002a: 248). Why indeed? Note that this principle is especially dubious for us here in light of the fact that we are explicitly discussing knowledge-*by-ordinary-standards*. Of course, Evans himself has made no mention of conceivability to back his pronouncement—because he has offered no backing at all. But similar problems would seem to arise if we instead ascribe to him thoughts about the mere possibility of one’s belief being false.

The need to locate *some* argument for the prohibition seems to me quite pressing, because the prohibition concerns the fairly technical, and also exceedingly murky, notion of a priori knowledge, and so can seem the kind of thing that cannot have much of a truly intuitive basis. Other ways out of the skeptical predicament before us seem, at least on the face of things, to bear real intuitive costs. But is it really “biting a bullet” to accept a priori knowledge_o of \sim BIV?

But it has been impressed on me that, for all my complaining about the murkiness of “a priori,” I really do have enough of a grasp of the notion to make some fairly clear rulings on whether or not bits of knowledge are a priori, and that a look at a bunch of fairly clear examples tends to promote the generalization that contingent facts are not known a priori. Don’t we understand the notion of a priori knowledge well enough to see that it is sensible to suppose that $2 + 3 = 5$ is (typically) known a priori, but not to hold that *The Red Sox won the 2004 World Series* is so known? And isn’t it the necessity of the former and the contingency of the latter that accounts for this difference?²⁹ Such an argument by examples may be what’s really behind the pronouncements of the likes of Evans. And it is a simple enough argument to sensibly suppose that the pronouncers might leave it unarticulated, if this is what’s behind their stance. Indeed, it might be a bit much to even call this an argument. Perhaps it’s better viewed as an attempt to bring to light (for stubborn resisters) a clear intuitive basis for Evans’s pronouncement: To show that, whatever murkiness there may be in the notion of the a priori, we still have enough of a grasp of it to make some quite

²⁹ This was urged to me by Brian Weatherson and Roger White, in a 2006 online discussion that took place on Weatherson’s blog, *Thoughts, Arguments and Rants*. The particular examples I use above were Weatherson’s.

clear judgments concerning it—judgments that form a pattern that points in the direction of the prohibition.

But I think that, while a look at the notion of the a priori reveals a sensible basis for drawing the line between what can and what cannot be known a priori much like the one Evans draws, it actually argues for drawing the line in a different place than Evans draws it—at a place that leaves the likes of *I am not a BIV* as a potential item of a priori knowledge, while ruling out as such the likes of *The Red Sox won the 2004 World Series*.

7.16. A Priori Knowledge: The Veins in the Marble

To find it strange (or worse) to suppose that *The Red Sox won the 2004 World Series* is known a priori while thinking that it is perfectly sensible to suppose that $2 + 3 = 5$ is so known, one need not be thinking of the a priori as some specially elevated status: You needn't think that our knowledge of the former is shown by such a ruling to be in any way less certain. This is important to adjudicating between Pryor-style dogmatism and my a priorism, because neither of us thinks of the a priori as such an elevated status—or at least I *think* this is a feature of my position that I share with Pryor. The warrant that one attains for one's perceptual beliefs in the way outlined by the basic conservative story is not supposed to be just some second-class variety of epistemic justification, which pales in comparison to the high-grade version of justification that is fitting only for certain highly privileged beliefs (or also for perceptual beliefs, but only after they have been properly based on the highly privileged beliefs). My conservatism, at least, runs very deep: The basic conservative story of how epistemic justification is generated for our perceptual beliefs is also the basic story of how we come to be justified in believing the likes of $2 + 3 = 5$. In both cases, we are justified in thinking that things are as they seem to us to be (absent good reason to think otherwise).³⁰ Or at least whatever difference in status there may be

³⁰ Here I take myself to be departing from the deeply anti-conservative attitude I find most emblematically expressed in the (disastrous in its implications) ninth paragraph of the Third Meditation, where Descartes distinguishes between “natural impulses” to belief and being able to intellectually “see” the truth of something by “the natural light.” Descartes is examining his long-standing belief in “things existing outside [him]” (Descartes 1996: 26) or “external objects” (Descartes 1971: 79), and has, in the previous paragraph, given this account of how this belief arose: “Nature has apparently taught me to think this” (1996: 26). Here is how Descartes draws the crucial distinction:

When I say “Nature taught me to think this,” all I mean is that a spontaneous impulse leads me to believe it, not that its truth has been revealed to me by some natural light. There is a big difference here. Whatever is revealed to me by the natural light—for example that from the fact that I am doubting it follows that I exist, and so on—cannot in any way be open to doubt. This is because there cannot be another faculty both as trustworthy as the natural light and also capable of showing me that such things are not true. But as for my natural impulses, I have often judged in the past that they were pushing me in the wrong direction when it was a question of choosing the good, and I do not see why I should place any greater confidence in them in other matters. (1996: 26–7)

isn't such as to ground a verdict that the former's justificatory status is initially problematic unless and until it can be shored up by properly grounding those beliefs on the initially favored beliefs of the latter type.

But even if the a priori is not some special, elevated status, there remains a potentially important question about the *source* of our knowledge (or potential knowledge) or justification, where we need not suppose anything about one source being more secure or elevated than another: Is a bit of knowledge, like our supposed knowledge, of our not being BIVs, something we *derive from* our experience of the world as opposed to something we *bring to* our experience of the world, the latter being the realm of the a priori? As Pryor points out (2000: 534), even our clearly a priori knowledge in some ways depends on our experience: As rationalists have long admitted, we might not have been able to entertain the relevant thoughts involved in a priori knowledge were it not for our experience of the world. Experience seems *in some sense* necessary for us to have any knowledge at all. Kant famously declared: "there can be no doubt that all our knowledge begins with experience." However, there are different ways that experience can be needed for bits of knowledge, and only some of these ways disqualify the knowledge from being a priori. Kant quickly assures us (in a way one may well find more confusing than assuring): "But though all our knowledge begins with experience, it does not follow that it all arises out of experience" (Kant 1929: 41). And that of course is just where the murkiness of the

His beliefs in external objects having fallen on the dark side of this crucial distinction, Descartes refuses to accept them, at least in his Meditations, unless and until they can be verified by things that are not just believed by "some blind impulse" (as he puts it in the twelfth paragraph), but that he can see to be true by the "natural light." Much is going on here, and I won't go into it all. But I should note that the "natural" belief that Descartes is speaking of here is not just that there are external (in some suitable sense) objects, but also that they in some seemingly murky and problematic way resemble his "ideas" of them: "But the chief question at this point concerns the ideas which I take to be derived from things existing outside me: what is my reason for thinking they resemble those things? Nature has apparently taught me to think this." Here, Descartes may well be saddling our "natural belief" with some extra baggage it doesn't actually carry, though issues of interpretation are tricky—not to mention issues about the actual nature of our relevant thoughts! And (especially for those of us with a history of actual error in matters of math and logic) much needs to be hashed out about how Descartes manages to individuate "faculties" so that believing what one sees to be true by the "natural light" presumably ends up with a pristine track record, as compared with the apparently very sorry record of following mere natural impulses to belief. But at the end of the day, the conservative, like me, thinks that Descartes is just fooling himself here: There is no great difference of the type Descartes thinks he has found. Ultimately, we are just going by how things seem to us. When appearances are in conflict, we give some up in favor of others, often with the aid of explanations of how and when we are apt to go wrong, and if all goes well (as it often enough does) we may in that way come to learn that some types of natural impulses tend to misfire. And on the other side, some of these would-be beliefs we find ourselves impelled toward fit together well with another, and survive and are bolstered by our best attempts to tidy up our view of the world, perhaps to the extent that we might start to think of them as having some altogether different and higher status than something that could just result from our best attempt to best manage appearances. But that way by which we come to "see" things to be true is in reality to arrive at a view of the world that is built on a suitably refined way of following our natural impulses toward belief, not to reach outside of what is indicated by our natural impulses and instead rely on a totally different (and pristinely trustworthy) source, as it seems that Descartes' "natural light" is supposed to be.

notion of the a priori resides: In discerning just which kinds of dependence upon experience disqualify knowledge from the title. And for me, at any rate, the best direction to turn in trying to draw the needed line, however imprecisely that is done, is by thinking in terms of a distinction between the information we learn by the particular way our experience of the world goes, versus what we are set up, prior to experience, to believe, just so long as we have enough, and varied enough, experience.

The great, old rationalist, Leibniz, memorably pointed toward such a distinction with his famous analogy of Hercules and the block of marble, where experience is represented by the “labor” the artist exerts on the marble:

I have also used the comparison with a block of veined marble, rather than a completely uniform block of marble, or an empty tablet, that is, what the philosophers call a *tabula rasa*. For if the soul were like these empty tablets, truths would be in us as the shape of Hercules is in a block of marble, when the marble is completely indifferent to receiving this shape or another. But if the stone had veins which marked out the shape of Hercules rather than other shapes, then that block would be more determined with respect to that shape and Hercules would be as though innate in it in some sense, even though some labor would be required for these veins to be exposed and polished into clarity by the removal of everything that prevents them from appearing. This is how ideas and truths are innate in us, as natural inclinations, dispositions, habits, or potentialities are. (Leibniz 1989: 294)

I like to modify Leibniz’s analogy a bit. Experience is represented by the “labor” of the artist. But we naturally imagine the artist as having her own plans for the block, which then may or may not be aligned with what turns out to be marked out in the veins of marble. I find it distracting from the important philosophical points of the example to think of experience as engaged in an effort to get a certain result. Better perhaps to think along the following lines, where the “artist” is viewed as working in an effectively aimless way. I have what seems to be a childhood memory (though I suspect it’s really something I only dreamed up myself in a “wouldn’t that be cool?” kind of way, but then later in childhood came to think of as something that really happened) of getting a children’s play sculpture kit as a Christmas gift. The kit consisted of a toy chisel and a toy hammer, and a block made of two different materials, each marble-like in appearance and visually indistinguishable from one another: A rock-solid material that the chisel could not possibly even scratch, in the shape of a bust of Abraham Lincoln (a three-dimensional version of the picture on the US penny), surrounded by a fairly brittle material that could be fairly easily chipped away by the chisel and hammer. We can now imagine even the quite haphazard “labor” of the five-year-old Keith, wildly wielding a child’s chisel and hammer with no discernible rhyme or reason, nonetheless reliably resulting in a bust of Lincoln emerging from our imagined block, at least given enough, and varied enough, “experience.” We do not have to suppose with Locke (Leibniz’s target) that our only choices for explaining how the mind might contain “innate principles” is to either suppose we are born already thinking the relevant thoughts (which is wrong)

or to merely suppose that we are born with some capacity or other to gain the needed concepts, given enough experience, and then have the relevant thoughts (which even Locke accepts). Rather, there can be *special* capacities, veins in the marble: Things that may depend on experience to emerge, but do not depend for their emergence on our experience following some narrow path, but that the mind is set up to (designed to, or as-it-were designed to, to taste) end up with, given just that it be subject to enough, and varied enough, experience.

I will now urge that viewed in the epistemically most relevant way, ~BIV is among the things we know_o in virtue of how the veins of our minds are configured, rather than something derived from the particular course of our experience. It is in epistemically relevant ways something we bring to, rather than derive from, our experience of the world. So our knowledge_o of ~BIV is, at least by the best sense I can make of the notion, a priori.

7.17. Deeply Contingent A Priori Knowledge and Radical Skeptical Hypotheses: Why Necessity Isn't Necessary

I start by explaining why I think the (deep) contingency of ~BIV is no bar to our knowing it a priori. The thought that there can be no a priori knowledge of deeply contingent facts isn't just some old philosophical prejudice: A rationale for the limitation can be found in the Leibniz-like notion (see Section 7.16) of what the mind brings to experience in virtue of "veins" it contains, as opposed to what it derives from its experience of the world.

We can appreciate the intuitive push behind the restriction by fancifully (and quite vaguely) supposing you are some kind of (quite limited, at least relative to Anselmian pictures) god, designing the processes by which humans come to hold their beliefs (thereby deciding which "veins" the blocks that are their minds will contain), seeking to make those beliefs reliably true, but working within some constraints. You can't be an occasionalist god, watching over your people in the various situations in which they find themselves and directly producing true beliefs in them, nor can you issue some kind of general divine directive like "Let them believe such-and-such if and only if (when and only when) it is true." Rather, you must build into them some belief-forming mechanisms (working within restraints we won't here spell out), and then leave the scene, not intervening further in people's cognitive lives. Well, what will these mechanisms have to be like? That matter looks very different for beliefs in necessary truths, as compared with beliefs in many contingent truths. If you want the people you design to reliably believe when it is true something like *There is a predator ahead of me* that from your general point of view may or may not be true (it will be true in some of the situations that your people find themselves in, but not others), it looks like you will need some mechanism that works at least somewhat like

perception, in that it is somehow “set off” by the presence of predators (or perhaps by something that will regularly enough accompany the presence of predators). Things are quite different when it comes to a necessary truth like $2 + 3 = 5$. On the downside, it doesn't look like any mechanism you might build into people could be “set off” by any state of affairs that makes it true that $2 + 3 = 5$, since that bit of arithmetic seems causally inert. But happily, it seems like you *don't need* a mechanism that works like that in order to design your people so they reliably form belief in that fact, because, since it is a necessary truth, it's the kind of thing that will be true wherever your people find themselves. So you don't need to worry about the mechanism producing the belief where it isn't true. You can just build into your people some a priori (as we can call it) mechanism that will reliably enough produce the belief, given just that they get enough, and varied enough, experience, without having to be “set off” by the fact that $2 + 3 = 5$. What this mechanism will end up looking like will depend on the exact nature of the constraints you're working under, as well as your divine whims, and perhaps it will end up looking in some ways much like the a posteriori mechanisms by which beliefs get “set off” by the facts that they're designed to detect. But perhaps not. The mechanisms that produce beliefs in the stable matters that you as a designer don't have to worry about being false can be quite different from the very basic type of mechanism that will be needed for reliably true beliefs about at least many contingent truths. These maximally stable necessary truths can be things that your believers won't need to learn from experience are the case, but are beliefs, or dispositions toward belief, that, with your help, they can *bring to* their experience, in virtue of being the results of a mechanism whose function is simply to produce the belief, as opposed to producing the belief where and only where it is correct. Which sounds quite a priori.

However, such intuitive grounds for the restriction on what kind of beliefs can be a priori suggest that we are not quite drawing the line in the right place in placing it between the necessary and the contingent (or between the necessary together with the shallowly contingent, on the one hand, and the deeply contingent on the other). For these considerations would seem to leave you as god free to use a priori mechanisms to get your people to believe *some* deeply contingent matters—namely, those that, though deeply contingent, are stably true in that they will be true wherever your people find themselves (or at least in enough of the situations in which they will be found). So, to use an example very relevant to an important form of skeptical argument that I haven't been addressing (but that is largely driven by supposed restrictions on what can be known a priori), suppose you know that some Principle of the Uniformity of Nature, whose exact content we don't need to specify here, and that we will simply label “U,” though deeply contingent, will be true in all (and therefore in enough of) the circumstances in which the people you design will find themselves. Well, then, to get them to form a reliably true belief in U, you don't need to give them some mechanism that will be “set off” by the Uniformity of Nature but that wouldn't be “set off” were Nature to fail to be Uniform in the relevant way (which may be a

242 HOW DO WE KNOW THAT WE'RE NOT BRAINS IN VATS?

very good thing, because concocting such a mechanism might be quite a tricky matter, depending of course on the exact content of U and the constraints you are working under). You can just build into them some a priori mechanism that will reliably enough lead them to form a belief in U . And the same goes for our beliefs that we're not BIVs.

Of course, an elegant design likely won't include distinct mechanisms working quite independently of one another, each of which exists to produce the belief that a particular global skeptical hypothesis is false—just as it won't contain special dispositions working quite independently of one another by which we come to believe various individual arithmetical facts. It would be deeply weird if Nature had endowed us with some special disposition, operating quite independently from dispositions to believe closely related facts, just to believe that \sim BIV (or one just for coming to believe that $2 + 3 = 5$). And one might suspect that a more realistic account of the way we actually do come to believe the likes of \sim BIV, and also other \sim Hs, has the potential to threaten the a priori status of our beliefs in the various \sim Hs. But that is a distinct worry (and one that we will soon be wrestling with) that does not involve the deep contingency of \sim BIV.

For the purpose of evaluating the deep contingency of \sim BIV itself as a ground for ruling we can't be a priori justified in believing it, we can quite fancifully imagine that it is the result of a quite simple belief-forming mechanism that is a priori in that its function is simply to produce the belief (rather than a posteriori in having the function of producing the belief where and only where it is true, by making our coming to hold it sensitive to how our experience unfolds). We can then compare that belief, together with a similarly formed belief in the also deeply contingent but stably true U , on the one hand, with a similarly formed belief in the both stably true and necessary fact that $2 + 3 = 5$. Taking a basic conservative perspective, we can suppose that all three beliefs are justified because they result innocently from natural belief-forming mechanisms, and then fit in well enough with our other beliefs that their presumptive justification is not overturned, but becomes part of our picture of the world that emerges from our best attempt to conservatively manage our appearances. Since we are supposing that the mechanisms by which these three beliefs are naturally formed are all of the type that work simply to form the belief in question (rather than to form it where and only where it is correct), but are highly reliable nonetheless, due to the stability (as I'm here using the term) of the truth of the beliefs that are their objects, the justification for believing each of these things looks a priori to me: These would all seem equally to be beliefs that in the relevant way we bring to, rather than derive from, our experience of the world. There doesn't seem to be anything *about the deep contingency* of the first two beliefs (in U and in \sim BIV) that would block them from constituting a priori knowledge, especially from a basic conservative perspective. And the justification for those beliefs, as well as for our arithmetical fact, looks to me a priori, too, in the best sense I can make of that term: It is justification that results from the beliefs seeming to us to be true through the

operation of an a priori mechanism—one designed simply to produce the belief in question, rather than to detect whether the content of the belief is true, and produce the belief if and only if it would be true.

7.18. A Less Fanciful Account of How We Come to Think We Are Not BIVs and the Dogmatist Account of How Such Beliefs Come to Be Justified

For my money, the more serious challenge to a prioritism comes not from the mere (deep) contingency of \sim BIV, but from more realistic thoughts about how we might come to believe that we're not BIVs. In Section 7.17, we fancifully supposed that we had some special belief-forming mechanism whose job it was to give each of us the belief that we are not a BIV. But one thing a somewhat more realistic account should do is explain how we might come to believe that any number of different radical skeptical hypotheses are false, without positing a different special mechanism for each. And the most plausible thought here seems to be that, through our experience of the world, we are led to form simple perceptual beliefs and then a picture of what the world is like whose content is at odds with those radical skeptical hypotheses. When those radical skeptical hypotheses then come to our attention, we are inclined to reject them, to the point of believing that they are false, because those hypotheses conflict with the highly justified picture of the world that we have arrived at. Supposing, plausibly enough, that our simple perceptual beliefs play a suitably foundational role in our coming to have the justified picture of the world that we end up with, the explanation for how we come to believe that radical skeptical hypotheses are false is ultimately that they conflict with our simple perceptual beliefs.

This gives us some reason to adopt the dogmatist position that our perceptual beliefs—the likes of O—serve for us as the *evidence* by which we come to justifiably believe that the skeptical hypotheses are false. After all, the perceptual beliefs are there first, and we come to disbelieve these Hs because of their conflict with those existing perceptual beliefs.

This case for dogmatism is enhanced if we further suppose that the resulting beliefs in \sim H depend for their justification on the justification of perceptual beliefs with which the radical Hs conflict. The alternative to dogmatism that I will adopt is that our justification for \sim BIV (and more generally, for the various \sim Hs) is *immediate*, not based on other beliefs at all. But a natural way—and the way that Pryor uses—to gloss “immediate justification” is to say it is justification a belief has that does not depend on the justification of other beliefs. So if we think that our justification for \sim BIV depends on our beliefs in the likes of O being justified, that gives us reason to think that we are basing our belief in \sim BIV on the likes of O. And, though I'm not certain of it, I am inclined to grant that on our story of how we come to believe that \sim BIV, this process yields a *justified* belief that \sim BIV only because the perceptual beliefs and the

picture of the world built up from them are themselves justified. It's hard to see how a belief is justified if it comes to be held because its opposite conflicts with a body of thoroughly unjustified beliefs.

7.19. Should We Ascribe a Basing Relation Here Without any Conscious Inference? Problems for Dogmatism (Again)

Should we then accept the dogmatist account of how our belief in \sim BIV comes to be justified? We are here approaching the tricky matter of determining when one belief is for us based on another, in such a way that the latter is serving as our evidence for the former.³¹ And on our common way of coming to believe that \sim BIV, I am not supposing that we go through anything like a conscious process of inferring that \sim BIV from any particular O, nor from any particular beliefs that were themselves based on such perceptual beliefs.

Of course, someone *could* talk or think themselves through the words of such an inference: "I have hands; BIVs don't have hands; ergo, I'm not a BIV." Such an "inference" will meet Moore's conditions for constituting a "proof" of its conclusion (Moore 1959a: 146), which we have already discussed (Section 3.2), at least where Moore's condition that one know the premise to be true is taken to require just knowledge_o, and in line with that, I will of course agree that one's possession of a genuine Moorean proof_o (as we might call it) does entail that one knows_o that one is not a BIV. But the inference here does not capture how we ordinarily come to believe or come to be justified in our belief that we're not BIVs, nor how we come to know_o it—nor, I think, given how we came to know_o the premise, how we even *could* come to know_o the conclusion. It is an exercise in constructing a Moorean proof_o for something we already believe and know_o. And the way we do ordinarily come to believe and know_o the conclusion of that "proof" is not through anything much like an inference at all, I am surmising, but rather by our rejecting the BIV hypothesis because it just seems so outlandishly (absurdly, ridiculously) false to us.

However, I don't think that the lack of anything like a conscious inference here automatically rules dogmatism out, because I don't think we should require that a believer goes through anything like a conscious process of inference from one belief (or group of beliefs) to another, nor even to consciously notice any relations of support that might hold between the beliefs, in order to rule that the believer is basing her second belief on the first in such a way that the first is serving as her evidence for the second. Perhaps, though she takes no notice of any relations of support, our

³¹ Greco considers the issue of whether our belief that we're not BIVs is based on inference to be a straightforward empirical matter that isn't philosophically important (2002: 561–2). For the reason about to emerge, I instead take the issue to have an important evaluative/philosophical aspect to it—though of course empirical results can be relevant and important to deciding it.

believer is appropriately *sensitive* in her holding of her second belief to the fact that it is supported by her first belief, where such sensitivity consists in such facts as that she wouldn't hold or continue to hold the second belief, or at least wouldn't continue to hold it to the degree that she does, were it not for the fact that it is supported by another belief she holds. One could be in that way sensitive to the support relations that hold among one's beliefs even where one gives no conscious thought to the relations, and yet it seems to me that epistemic justification might very well be transmitted from one belief to another in virtue of such sensitivity to relations of support.

But if *that's* how we are going to reach the determination that we base our belief in \sim BIV on our beliefs in the likes of O—if it's not because we engage in anything much like a conscious process of inference, but rather because we are just appropriately sensitive to the support that our belief in \sim BIV receives from our beliefs in the likes of O—then there had better be support relations for us to be appropriately sensitive to. But if, as I am supposing (see Section 7.14), our “Problems for Dogmatism” are real, then there are no relations of support for us to be sensitive to. Though our beliefs in the likes of O *entail* \sim BIV, given how we come to hold these beliefs in the likes of O, they are undermined for use as evidence for the likes of \sim BIV, and do not support such “conclusions.” (For the same reason, in our baseball report case from Section 7.14, unless she actually goes through some process like making the suspicious inference, we should not attribute reliance on the transparently bad “track record argument” to a normal person, quite reasonably (to all appearances) believing both that the Cubs won and that the radio report she heard was reliable, even though the premises of that argument *would* support the conclusion that the report is reliable if the subject had come to hold all those premises in a way that did not render them undermined for the purpose of supporting that conclusion—and even if she might not be able to come up with anything very impressive if she were challenged on why she does believe the report is reliable.)

7.20. An A Prioritist Account

So I say: Once we see that, for reasons given in Section 7.17, \sim BIV is not ruled out by considerations of its modal status as an item of a priori knowledge, we are free to take it as an item of immediate justification—and we should take advantage of that opportunity. No longer constrained to seeing that belief as being based on empirical evidence, we can, and should, take its justification to be not based on evidence at all, along the lines of the following account.

Our disposition to believe that \sim BIV is something we bring to, rather than derive from, experience, though our beliefs in the likes of O play a key role in the process by which we come to believe \sim BIV: The way that \sim BIV is encoded in the veins of our minds is that we are disposed to form simple perceptual beliefs whose content puts them at odds with BIV. We are disposed to experience the world in accord with what

is sometimes called the “Real World Hypothesis”—though it then doesn't really function as a “hypothesis” for us, which gets confirmed or justified by its ability to explain the empirical evidence we gather; so let us call it the “Real World View” (RWV). When we form perceptual beliefs with real world content, and then hear about and reject the BIV hypothesis because it conflicts with the RWV picture of the world those perceptual beliefs have led us to form, that is how our disposition to believe \sim BIV is activated. (RWV is then like the bust of Lincoln in our analogy. The BIV hypothesis is like some alternative to a bust of Lincoln—say, a bust of President Andrew Johnson—that the block of marble does not contain in virtue of its containing a bust of Lincoln.)

As I'm inclined to admit (as we have seen), we would not be justified in believing \sim BIV if our beliefs in (the likes of) O were not justified. However, that should not be taken to show that O is functioning for us as *evidence* for \sim BIV, for *that* weak kind of “would not be justified were it not for our justification for the other one” dependence is *mutual* between (the likes of) O and \sim BIV. We are only justified in believing O (and its cousins), with its robust real world content, because of the *a priori* (as I am inclined to classify it) justification we have for \sim BIV.

This dependence is perhaps best seen by considering what happens when our justification for \sim BIV is lost. Though I hold that one's belief in \sim BIV is not based on empirical evidence, that justification can be lost through empirical evidence. One's experience could go in such a way that it becomes a very serious possibility that one is a BIV and one is no longer justified in believing that one is not. Suppose, for example, that you (at least seem to) learn that super-advanced space aliens have gone into orbit around the Earth, and, as they have told us they would do, they have been “snatching” some people's brains from time to time for their envatment experiments. (“Sorry, but...”) They do so in a fast, painless, super-advanced way such that their victims can't tell that they've been envatted. To these victims, their life seems to be going on as before. (The computer has pre-scanned their brain, and gives it a train of continuing experiences that will make good sense to the victim as a continuation of their life.) As a consolation to the loved ones of the victims, the aliens leave near each victim's debrained, lifeless body a TV which shows from the victim's point of view what's happening in the victim's virtual world. Now, suppose that the aliens seem to be closing in on you: They've been working your block, and several of your neighbors' brains have been snatched and envatted, their grieving families gathering around the TV sets showing what's happening in their loved ones' virtual worlds. By a certain point in this closing-in process, it will become a *very* serious possibility to you that *you* have been envatted, and you will no longer be justified in believing that you are not a BIV. (We can suppose that now the aliens have (seemed to have) snatched a member or two of your family that lives with you. In an interesting twist, we can suppose that while you (seem to) watch the TV that shows what's virtually happening with your daughter, you (seem to) discover that in her virtual world, she has just discovered that you have been envatted, and is watching the TV showing

your virtual world. We can suppose that you have now come to (seem to) be the only person on your block who has *not* been snatched, but you (seem to) see from the various TVs in your neighbors' apartments that *they* each seem to themselves now to be the only person on the block yet to be snatched.) You are still wired to form perceptual beliefs with real world content, and so you still, in a relevant, good sense, *seem* to yourself to be experiencing a real world (and we can suppose you in fact are the one who has not yet been snatched, and so this appearance is veridical), but it seems that as you become unjustified in believing that you are not a BIV, you cease to be epistemically justified in holding the perceptual beliefs, with their real world content, that you still have an inclination toward.

In discussing whether one's belief in \sim BIV is based on (the likes of) O in such a way that the latter serves as our evidence for the former, we are not asking about causal dependence—whether one's beliefs in the latter play a causal role in one coming to believe the former. We are asking about *rational* dependence, or what Pryor calls “epistemic priority” (Pryor 2000: 524–6). And about *that* issue, it seems the verdict to reach on the basic conservative, direct realist (as I think it can be well called) picture we have been pursuing, neither of these beliefs is evidentially based on the other. They are each such that we would not be justified in believing it were it not for our justification for believing (at least the likes of) the other, but that's not because of any evidential basing, but because our coming to have our justification for each is realized in the process by which we come to form our perceptual beliefs in the likes of O, in such a way that the justification for either would tend to stand or (more to the point) fall with that of (the likes of) the other.

7.21. Dogmatism vs. A Prioritism on Two Questions

As we have seen, I differ from dogmatism over both the question of whether our knowledge_o of \sim BIV is a priori, and whether the likes of O serve as the evidence by which we come to know_o (and come to justifiably believe) that we're not BIVs. However, the one disagreement has a different status from that of the other. Though, as we are about to discuss, it is the first disagreement that drives different responses to Split Cases AI, with the dogmatist and I denying different premises of that argument (where its instances of “know(s)” are understood to be indexed to ordinary epistemic standards), and though it is from my answer to the first question that my “a prioritism” derives its name, it's actually the second disagreement that is the more substantial one.

My contentions that our beliefs in the likes of O *cannot* properly be used as evidence for the likes of \sim BIV, and therefore that basing our belief in \sim BIV on the likes of O in such a way that the latter serves as our evidence for the former *is* not the way we come to know_o and justifiably believe that \sim BIV, constitutes an important and substantial epistemological difference with dogmatism.

By contrast, I don't take the meaning of the philosophical term "a priori" to be settled enough to securely put the dogmatist and me on opposite sides of the question of whether our knowledge_o of \sim BIV is a priori. That our knowledge_o of \sim BIV is a priori in the way I use the term is an important part of my position, and in large part explains why I feel our belief that \sim BIV doesn't have to be evidentially based on the likes of O in order to be known_o and epistemically justified. But that important part of my position does not depend on my preferred way of using "a priori" being decidedly the correct way.

My attitude toward our having a priori knowledge of \sim BIV is very much like that displayed more than forty years ago by Alvin Plantinga toward the question of whether we have any a priori knowledge of contingent truths (Plantinga 1974: 6–9). (I should note that Plantinga doesn't include anything like the denials of radical skeptical hypotheses among his examples of contingent truths we know a priori.) Plantinga complains that the distinction between a priori and a posteriori knowledge is "clouded in obscurity" (1974: 6), but, based on "the rough and intuitive understanding" he thinks we have of the terms involved, he lists some examples of contingent truths he is inclined to rule are known a priori, and concludes, "It is fair to say, therefore, that I probably know some contingent truths a priori" (1974: 8).³² Well, that's how I feel about our knowledge that we're not BIVs. On the best sense I can make of the notion, I think that knowledge is best classified as a priori. But someone could understand my position on the issues perfectly well and think that my position should be classified as one on which our knowledge_o of \sim BIV is not a priori. I would think they are not using "a priori" in the *best* way, but that their use of the term is not decidedly wrong.

7.22. An Evaluation of Split Cases AI

Recall, then:

Split Cases AI (AI-SC)

- 1a. I don't know a priori that not-H.
 - 1e. I don't know through experience that not-H.
 2. If I don't know that not-H, then I don't know that O.
- So, C. I don't know that O.

³² A similar attitude is displayed at Manley (2007: 409–14), where Manley gives examples in which his safety condition strains "traditional notions of apriority" (2007: 409), and seems inclined to rule that the examples he gives should be counted as ones that feature a priori justification, but adds in a note: "I suspect it is vague whether the pieces of knowledge in question are a priori. Terms of art are not after all immune to vagueness, especially those with so long and varied a history as 'a priori.' I have no objection to putting 'a priori' to a precise use, as long as one acknowledges the technical nature of the term" (2007: 414, n. 31).

And recall our resolution, however exactly we understand “a priori” and “through experience,” to coordinate our construals of those two terms so that it will be true that all knowledge one has must be either one or the other.

This argument can seem an improvement on AI since, in breaking AI's first premise into two cases, it can utilize the ever-popular restriction against our having a priori knowledge of deeply contingent facts to support 1a, and then, to motivate 1e, it can tap into strong intuitions to the effect that the things we learn about the world through experience get “undermined” if we try to use them as evidence for \sim H.

But we must be alert to just how slippery this argument is. For reasons we have already seen, when the occurrences of “know” in it are understood as indexed or ordinary epistemic standards, I think this argument is best handled by denying 1a. This is based on the best sense I can make of the notion of a priori knowledge. However, so far as I can see, the notion of a priori knowledge (and then also the coordinated notion of knowledge gained “through experience”) is fairly malleable, and those terms can be used in acceptable ways such that (with no real change in my views) I will instead think that it's 1e that is mistaken. Though it is not the judgment I feel the most confident about, I have agreed that our justification for rejecting the BIV hypothesis depends on our being justified in accepting the beliefs that, by conflicting with the BIV hypothesis, lead us to believe that \sim BIV. And though I don't think this is the *best* way to go, so far as I can see, it is *allowable* to use the notion of a priori knowledge so that our knowledge of \sim BIV is by definition not a priori if our justification for believing \sim BIV depends on our justification for holding beliefs we have gained (in the relevant sense) “through experience,” whether or not one is really basing one's belief in \sim BIV on those beliefs gained through experience. This gets messy when applied to my view, for, first, I don't hold that our justification for \sim BIV depends on our justification for any particular perceptual beliefs (as usually happens when some belief or small cluster of beliefs function as evidence for others), but rather on enough of our perceptual beliefs, taken as a group, being justified. And, second, even taking our perceptual beliefs as a group, the dependence here is *mutual*, at least where “dependence” is understood counterfactually: Our justification both for the various Os we believe and for \sim BIV are such that we would not have it if we did not have the justification we have for the other. But I can see nothing to really block someone resolving their use of the relevant terminology in such a way that any such dependence, mutual or not, and involving any real basing or not, of the justification for \sim BIV on the justification for any beliefs, or groups thereof, gained “through experience,” rules \sim BIV out from being a piece of a priori knowledge.³³

³³ Alternatively, they can appeal to a way of understanding “dependence” that focuses on the temporal or causal order in which the beliefs are formed, so that the dependence involved here is one-way, rather than mutual. But to my thinking, this is to lose sight of the fact that we seek an account of rational dependence, not an account of causal dependence in the way the relevant beliefs are formed.

I would not object (at least not strenuously) to such uses of our key terms. I would then just reject 1e: In such a weak sense (*too* weak to be optimal, by my lights), our knowledge_o of and our justification for \sim BIV are gained “through experience,” and are not (in the correlatively overly strong sense) a priori. These notions are too malleable for it to be wise to really lock into one rather than the other way (the denying 1a vs. the denying 1e way) of resisting AI-SC.

But what we *definitely* should *not* do is grasp on to the old “No a priori knowledge of deeply contingent truths” mantra in such a way that we accept it for just any allowable way of drawing the distinction between a priori knowledge and knowledge that we gain through experience. (This is especially true for me, since, for reasons we've seen, I think the mantra is false on the *best* way of drawing the distinction.) I think that the skeptic's best legitimate strategy for motivating 1a and 1e together is just to appeal to the powerful intuitive support (which we discussed in Section 2.9) for good old 1 (the first premise of original AI) which, on any way of working out the meaning of our key terms, implies both 1a and 1e: Which is to say I think the skeptic does not really gain any legitimate added power for her case by moving from AI to AI-SC, but rather, in effect, takes illicit shelter in the slipperiness of the key notions involved in this slightly souped-up argument. But it is up to the skeptic who thinks differently to tell us how she is drawing the distinction between a priori knowledge and knowledge gained through experience and why 1a and 1e are both compelling, powerful, or at least very plausible, given her way of drawing that line. If we keep our wits about us, no way by which she might make the distinction should by itself have the power to scare us off of our basic conservative, easy account of how we are justified in believing (and how we know_o) that radical skeptical hypotheses are false. Supposing we have escaped bold skepticism based on the original AI, it seems to me that the skeptic's best bet for regrouping and trapping us with this slightly souped-up version of the argument is by hoping we will unthinkingly accept her mantra, so she can then, with acceptance of 1a anchored firmly in place, push for a way of drawing the a priori/through experience distinction that is favorable to the truth of 1e. Let's not fall for that.

7.23. How Do We Know that We're Not Brains in Vats?

So, now setting aside the murky and distracting matter of whether this knowledge is a priori, how do we know_o that we're not BIVs? How could we *possibly* know something so exotic like that, in any way? In short, we are inclined to believe it, and we do believe it, and we have no good reason to think we are wrong to so believe: When we start playing the various things we're inclined to believe off one another, to arrive at a picture of the world that makes the best sense of as many of those things as possible, it is not among those that get thrown out. And we're right about the matter

HOW DO WE KNOW THAT WE'RE NOT BRAINS IN VATS? 251

(I claim, perhaps provocatively, perhaps somewhat naughtily): We are not BIVs. Nor were we (in the relevant sense) close to being wrong about this issue. In shorter, we know_o that we're not BIVs in the same basic way that we know just about anything. It's no big thing. It's just one of those things that we know_o that we can seem not to know at all—at least when the issue is viewed from certain angles. The trick is explaining why we can seem not to know it.