ESSAYS ON SKEPTICISM ABOUT EPISTEMIC REASON

by

Timothy Willenken

B.A., Yale University, 2003

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Timothy Willenken

It was defended on

April 11, 2011

and approved by

John McDowell, Distinguished University Professor of Philosophy, University of Pittsburgh

Cian Dorr, Lecturer in Philosophy, Oxford University

Anil Gupta, Distinguished Professor of Philosophy, University of Pittsburgh

Karl Schafer, Assistant Professor of Philosophy, University of Pittsburgh

Dissertation Advisor: Kieran Setiya, Associate Professor of Philosophy, University of

Pittsburgh

ESSAYS ON SKEPTICISM ABOUT EPISTEMIC REASON

Timothy Willenken, Ph.D

University of Pittsburgh, 2011

Most of us believe that induction and perception have some normative status that counter-induction and crystal gazing lack: the former are correct, but the latter are not. How are such beliefs about rationality justified? My dissertation examines two skeptical arguments that contend the answer is: they're not.

The first skeptical worry centers on *circularity*. The only defense I can give for the claim that induction will mostly lead me to true beliefs will invoke induction – it will point out that induction has been reliable in the past and thus conclude (via inductive inference) that induction will be reliable in the future. Much the same applies to perception: I can give a story about why I expect it to be reliable, but only by citing perceptual beliefs. These defenses seem worryingly circular.

Non-skeptical responses to this puzzle fall into two camps: Mooreans embrace the circular defenses of perception and induction; rationalists say that justification to believe that perception and induction are reliable is apriori. I defend Moorean responses to skepticism: the most plausible accounts of why the aforementioned reasoning is viciously circular fail. In addition, I argue that rationalism—while perhaps true—is insufficient to deflect the skeptical worry. It turns out that even rationalists need to embrace Moorean circular reasoning.

The second skeptical worry focuses on the *etiology* of our faculties of reason. There is some causal story about why I am inclined to engage in certain patterns of normative reasoning: roughly, evolution by natural selection. Selection pressures favored norms that helped our ancestors find food and show off to potential mates. A puzzle arises because correctness does not appear well-positioned to provide an adaptive edge. The correct ways of reasoning about normative matters might have aided survival, but only as a fortuitous side effect - so getting it right would be a fluke.

I show that this puzzle yields a serious skeptical worry. We ought to doubt that we are trustworthy normative reasoners unless there is an explanatory connection between the normative facts and our faculties for normative reasoning.

TABLE OF CONTENTS

PREFACE

I would like to thank my committee—most especially Cian Dorr and Kieran Setiya—for the great compliment of high expectations. I am not sure if I have met these expectations. But I am sure that they helped me to learn, at least, how to give a task my all. I hope that one day I will be able to pass this favor on.

1.0 INTRODUCTION: HOW CAN I KNOW WHETHER I AM REASONING CORRECTLY?

People can and do form beliefs in myriad ways: induction, gazing into crystal balls, perception, consulting sacred texts, and so on. In addition to this actual diversity, we can imagine possible thinkers who employ still more exotic approaches: counter-induction, inference to the funniest explanation, or whatever. Not all epistemic differences are so startling: among practitioners of induction, some embrace more liberal versions of induction (becoming confident that observed patterns will continue based on relatively little evidence) whereas others are more conservative (remaining agnostic about the future until the evidence of a pattern is overwhelming). We ordinarily think that some of these ways of conforming one's beliefs to the evidence are correct whereas others are not. When we condemn the pathological rules, our thought isn't just that we think such rules are practically ill-advised, as when we condemn someone for (say) not saving any money. Rather, we think they involve going wrong about what is evidence for what: we think that other people are following the wrong **evidential norms**.

What separates the correct evidential norms from the incorrect ones? In some cases, the answer is easy. I have good inductive justification to believe that *The New York Times* is pretty reliable and that *Weekly World News* is quite unreliable, so the norm of believing what I read in the *Times* is correct while the norm of believing what I read in the *Weekly World News* is not. However, we can't account for the correctness of every norm in this way. At least some norms must be **basic**, where N is a basic norm just in case it is appropriate for me to reason in accordance with N absent a belief justified by *another* norm that reasoning in accord with N is reliable. So the really difficult challenge is to say what separates the correct basic evidential norms from the incorrect ones. We can separate this puzzle into two questions: first, what separates the correct basic evidential norms from the pathological ones? Second, why is it reasonable for me to believe that *my* basic evidential norms are correct?

1

A tempting answer to the second question is: it isn't. The answer is tempting because of two skeptical arguments that purport to show that it is unreasonable to believe that your evidential norms are correct. These arguments allege that our beliefs about evidential norms are no better than blind stabs in the dark.

Setting out and responding to these two skeptical arguments will be the main task of this dissertation. Here I will give a brief sketch of the skeptical arguments and of the overall shape of my reply. Setting out these problems with care takes some work, though, so what I say here will necessarily be rough – for the best statements of the skeptical problems, see later chapters. My focus will be on the question of how it could be rational to believe that my norms are the correct ones, but, as we will see, this requires making some claims about just what kind of property correctness is.

1.1   FIRST WORRY: CIRCULARITY

Suppose someone asked you to give a defense of induction; to explain why you think it reasonable to believe the future will resemble the past. It's a familiar point that the best we can offer is a circular reply: pointing out that, in the past, the future has always resembled the past and thus concluding (via inductive inference) that in the future, the future will resemble the past. Much the same applies to perception: if pressed on why you trust your perceptual beliefs, you can give a reply – but that reply will have to invoke beliefs based on perception. These replies look viciously circular. If they are, we are left without any means of justifying our beliefs that perception and induction are better routes to the truth than just picking claims to believe out of a hat. We can state the argument for inductive skepticism a bit more carefully

> **Premise:** Any justification I have to believe that the future will resemble the past must depend (at least in part) on induction
> **Premise:** I can form justified beliefs using induction only if I *already* have justification to believe that the future will resemble the past
>
> **So:** I do not have justification to believe that the future will resemble the past

We can generalize this form to yield a skeptical argument concerning any basic evidential norm N:

> **Premise:** Any justification I have to believe that N is reliable must come from reasoning in accord with N
> **Premise:** I can form justified beliefs using N only if I *already* have justification to believe that reasoning in accord with N is reliable
> ___
> **So:** I do not have justification to believe that reasoning in accord with N is reliable

Since N is basic, the first premise looks true – if reasoning in accord with N were appropriate because of a belief that reasoning in accord with N is reliable justified by some other norm N' then, by definition, N would not be basic. The second premise looks like a plausible anti-circularity constraint. But the conclusion—that none of us is justified in counting his basic evidential norms as any better than guessing—is devastating. How can we reply?

One strategy for reply is **rationalism**: rejecting the first premise of each argument on the grounds that I have apriori justification to believe my basic norms are reliable (including, e.g. apriori justification to believe that the future will resemble the past). A second strategy is **Mooreanism**: rejecting the second premise of each argument and saying that I can, for instance, acquire a justified belief that the future will resemble the past for the first time using induction (this position gets its name from G. E. Moore's infamous "proof" of an external world). Moorean inferences look viciously circular, though, so we seem stuck with two unappealing options: embracing implausible apriori justification or condoning some fishy-looking circular reasoning.

1.2 WHAT I WILL SAY ABOUT THE FIRST WORRY

I agree that Moorean responses to skepticism look fishy, but I will defend them nonetheless. One way to defend Mooreanism—a strategy that I will *not* adopt—is to say that it is less bad than skepticism or rationalism; since we are forced to choose, we ought to hold our noses and

become Mooreans. Perhaps this line of reasoning is compelling, but fortunately matters are not so bleak. There is no good reason to condemn Moorean responses to skepticism. I will begin my defense of Mooreanism with a question: what, exactly, is circular about Moorean responses to skepticism? After all, Moore's "proof" and the inductive defense of induction do not have conclusions that appear among their premises. The standard account of what is wrong with Moorean reasoning is that it violates a principle called **conservatism:** the claim that someone can form justified beliefs via some reasoning only if he *already* has justification to believe that all skeptical underminers for that reasoning are false.

Most fans of Mooreanism reject conservatism. I don't. Instead, I argue that conservatism is ambiguous because it employs the ambiguous notion of *justification to believe*. Philosophers normally understand *S has justification to believe that p* in two ways:

1. S has adequate evidence that bears on whether p

2. There is a sound, undefeated route of reasoning from S's evidence to the conclusion that p

These two understandings are not equivalent. They come apart quite dramatically in cases where S has lots of evidence in favor of p but also has lots of evidence that *he cannot figure out whether p on the basis of his evidence*: in such cases S has justification to believe that p in the first but not the second sense. The two understandings of *justification to believe that p* come apart because someone can reasonably conclude that p only when two kinds of evidence are in order: evidence that p (call this **object-directed evidence**) and evidence that he can figure out whether p (call this **reasoning-directed evidence**). It is ambiguous whether conservatism uses the term *justification to believe* in a strong sense that requires having both kinds of evidence in order or a weak one that requires only the first kind of evidence. In chapter 1 I argue that plugging either the stronger or weaker of these two senses of *justification to believe* into conservatism yields a claim that does not make trouble for Moorean responses to skepticism. If we plug the stronger sense in, conservatism is false – it badly misfires in humdrum cases. If we plug the weaker sense in, conservatism may well be true, but it doesn't condemn Moorean responses to skepticism. Either way, Mooreans are out of the woods.

What about the rationalist anti-skeptical strategy? In chapter 2, I will lay out a new challenge for fans of rationalist responses to skepticism. A common worry about rationalism is that it is mysterious or implausible – it's not totally obvious where all of this apriori justification is supposed to come from. I'll set that worry aside. Instead, I'll focus on the worry that rationalist responses to skepticism end up embracing the same kinds of circular reasoning as Moorean ones. In particular, I will focus on a skeptical scenario tailor-made to cause trouble for rationalism:

> I am a recently envatted brain and a neuroscientist is manipulating my faculties of reason in ways that make it seem to me that I have apriori justification to believe false, deeply contingent propositions for which I do not, in fact, have apriori justification.

If a rationalist responds to this scenario by saying "it's just apriori that I'm not in that kind of situation" he seems to be begging the question against skeptics in *just the same way* that Mooreans do. In chapter 2, I argue that this appearance is correct: there is no plausible anti-circularity constraint that rules out Moorean reasoning but does not rule out the rationalist reply to the above skeptical hypothesis. As a result, rationalists and Mooreans are on all fours when it comes to making circular arguments. Does that mean rationalism is *false*? No. But it does mean that worries about circularity provide no reason to prefer rationalist responses to skepticism over Moorean ones. And this, in turn, entails the surprising result that the truth of rationalism— apriori justification to believe that induction and perception and so on are reliable—isn't, by itself, an adequate response to skepticism about epistemic rationality. If circular arguments (in the sense in which Moorean arguments are circular) are no good, skeptics win *even if rationalism is true*. If circular arguments are fine, Moorean responses to skepticism are adequate. All that apriori justification doesn't make the need to embrace Moorean reasoning any less pressing.

1.3 SECOND WORRY: THE ETIOLOGY OF NORMATIVE JUDGMENTS

There is some causal story about why I embrace the evidential norms that I do: roughly, evolution by natural selection. Consider **non-naturalistic realism** about norms: the view that our attitudes do not explain why some epistemic norms are correct (that's realism) and correctness is not identical to any natural properties (that's non-naturalism). Non-naturalistic realists face a puzzle about why anyone should think that her norms are correct. Selection pressures favored norms that helped our ancestors find food and show off to potential mates. A puzzle arises because correctness—if it's a non-natural property of norms—provides no adaptive edge. The correct ways of reasoning about normative matters might have aided survival, but only as a fortuitous side effect. As a result, getting it right would be a fluke. That seems a compelling reason to doubt that I am getting it right. I ought to think that, unless I've fortuitously stumbled upon a winning ticket in the normative lottery, my evidential norms are not the correct ones.

1.4 WHAT I WILL SAY ABOUT THE SECOND WORRY

The second worry—rooted in the etiology of our faculties of judgment—is serious. Many standard anti-skeptical strategies (in particular, Mooreanism and rationalism) cannot deflect it. The only way to avoid skepticism is to reject the assumption that gets the skeptical worry off the ground. We ought to trust our normative judgments only if there is some explanatory connection between the normative facts and our faculties of normative judgment: our faculties must explain why the norms are what they are or the norms must explain why our faculties are what they are. Non-naturalistic realism does not allow either direction of explanation—the problem isn't that non-natural facts are causally inert, but rather that even on views wherein they have causal powers they aren't well positioned to explain why I reason in accord with one norm rather than another—so it leads to skepticism.

Three notes about this skeptical strategy: first, though naturalists and constructivists/anti-realists (people who think our proclivities to make certain judgments explain why the normative facts are what they are) can resist this skeptical attack, they can do so only by embracing a kind of Moorean inference. So Mooreanism isn't sufficient to block this skeptical worry, but it is necessary. Second, nothing in the skeptical argument turns on the details of evolution by natural selection. The attack is utterly apriori. It turns on two apparently innocuous claims about evidential support:

1. It is rational to believe the best explanation of your evidence
2. If [someone can form a justified belief that p on the basis of process R] then [the output of R is evidence that p]

Third, the skeptical worry threatens to generalize to beliefs about any domain wherein there is no explanatory connection between the facts in that domain, on the one hand, and our beliefs about the domain, on the other. Our beliefs about logic, mathematics, and metaphysics all appear to fall within this category. Though I do not pursue these extensions of this skeptical worry in what follows, they mean that the stakes in understanding this skeptical mode are high indeed.

1.5   WHAT ALL OF THIS, TAKEN TOGETHER, SHOWS

The need to respond to the first skeptical attack does not put any constraints on the nature of correctness, but does require that we embrace Moorean reasoning. That's not bad news, though, since the standard case against Moorean reasoning is no good. The second attack does require us to adopt a certain story about the nature of correctness, on pain of thoroughgoing normative skepticism: there must be an explanatory connection between our normative judgments and the normative facts. It *also* requires us to embrace Moorean reasoning. Moreover, rationalism—whatever its merits—is insufficient to deflect either skeptical attack.

7

Before commencing our argument in detail, though, a note on the argumentative strategy of all three chapters contained herein: each discussion of skepticism will, at a key juncture, invoke reasoning-directed defeat. That is, a skeptical move will involve attacking justification to believe that p not by producing evidence that not-p, or by undermining one's evidence for p, but rather by giving reason to doubt that you can figure out whether p. A number of writers have, of late, begun to emphasize the hitherto ignored topic of reasoning-directed evidence (often called—misleadingly—"higher-order evidence"): it figures prominently in recent discussions of the epistemic significance of disagreement and the shortcomings of Bayesianism as a guide to rational belief. One upshot of this dissertation is that these recent discussions reveal just the tip of an iceberg: reasoning-directed defeat is at the heart of many, perhaps even all, modes of skeptical attack. The relative dearth of literature on reasoning-directed evidence is thus no small oversight. In what follows I will offer some guidance about the rational significance of reasoning-directed evidence. Mostly, though, this dissertation serves to show that puzzles about reasoning-directed evidence are central to epistemology. Questions about how to respond to evidence of our own cognitive limitations lie at the heart of some of the deepest and most pressing skeptical worries.

2.0   MOOREAN RESPONSES TO SKEPTICISM: A DEFENSE

Many philosophers believe that it is hard, or even impossible, to show that some notorious skeptical hypotheses are false. But according to many followers of G. E. Moore, refuting these skeptical hypotheses is *very easy*. Consider:

> **Enabling condition[1]: (**Moore-1) [an experience as of a hand]
> **So: (Moore-2)** I have a hand
>
> ---
>
> **Therefore: (**Moore-3) I am not a handless brain in a vat
>
> **Premise:** (Induction-1) In the past, observed regularities have tended to continue into the future
> **So:** (Induction-2) From now on, observed regularities will tend to continue into the future
>
> ---
>
> **Therefore:** (Induction-3) Induction will continue to be reliable in the future

Can we stop worrying about these skeptical hypotheses?

Few of us think it's so easy. These bits of Moore-inspired reasoning look fishy indeed. It turns out to be difficult, however, to say just what is wrong with them. None of the above bits of reasoning is *tautological*: their conclusions do not appear among their premises. Denying the premise Induction-1, or denying that we are ever in the enabling condition Moore-1, is totally implausible. Moreover, in each case the premise or enabling condition rationally supports an intermediate result, and the intermediate result entails the conclusion. Of course, someone could generate a problem for these arguments by denying that we are entitled to believe the

---

[1] I say "enabling condition" rather than "premise" since having some perceptual experience cannot itself be a premise in an argument (though, of course, the proposition *I am having a perceptual experience as of a hand* could be). By *enabling condition* I simply mean a state that licenses the formation of another belief, here via the use of perception.

intermediate result—for instance, that I have a hand—but that seems like an awfully drastic measure. We would like a diagnosis of what is fishy about these arguments that does not commit us to skepticism about the existence of nearby mid-sized objects. Finally, we could generate a problem for these arguments by denying **closure**—the thesis that if someone has justification to believe that p, and p entails q, and he forms a belief that q based on p, then his belief that q must be justified—and then denying that anyone who has justification to believe the intermediate result of each argument must have justification to believe the conclusion. But denying closure is a radical step indeed.[2] So far none of the possible diagnoses of what's wrong with these arguments looks remotely appealing. Can't we do better?

Yes. A better diagnosis rests on the observation that in both cases the conclusion of the bit of reasoning negates a skeptical hypothesis that removes the support lent by the premises or enabling condition to the intermediate result. For instance, the negation of Moore-3 removes the support lent by a perceptual experience as of a hand for the proposition *I have a hand*. The plausible diagnosis rests on a principle called **conservatism**: roughly, the claim that someone can acquire a justified belief that p on the basis of E only if he already has justification to believe that all of the skeptical hypotheses that undermine the support lent by E to p are false. Conservatives say that the problem with each of the above bits of reasoning is that nobody can ever use them to acquire justification to believe their conclusions. Why not? Well, someone can acquire justification to believe Induction-2 on the basis of Induction-1 only if she *already* has justification to believe that all of the skeptical hypotheses that remove support for inductive inferences—including the negation of Induction-3—are false. So someone is in a position to perform the reasoning described in Induction 1-3 only if she already has justification to believe Induction-3. The same diagnosis applies to Moore 1-3. So these bits of reasoning—while not tautological—are epistemically useless. Nobody can use them to acquire justification to believe anything new.[3]

Part of the appeal of the conservative diagnosis is that in everyday cases—as opposed to cases involving notorious skeptical hypothesis—Moorean reasoning looks obviously futile. Consider:

---

[2] Any adequate argument in favor of closure would take us far from the main thrust of this paper. For a full defense of closure, see Hawthorne.

[3] The original statement of this diagnosis is Wright (1985); further refinements appear in later Wright papers, as well as White.

Premise: (Table-1) This table looks red
So: (Table-2) This table is red

Therefore: (Table-3) This is not a white table under red lights

Premise: (Election-1) Someone just checked a box on a piece of paper and slid it into a slot
So: (Election-2) Someone just voted

Therefore: (Election-3) An election is happening, rather than a rehearsal

These look like terrible bits of reasoning. It should not be so easy to acquire justified beliefs in Table-3 or Election-3; if these arguments are not epistemically useless they generate odious "easy justification" for their conclusions.[4] Yet, as before, it can be tricky to say what is wrong with each bit of reasoning: neither is tautological, and each step looks individually unobjectionable. Conservatism is well-positioned to explain the source of our unease: in both cases, the conclusion undermines the support lent by the premise to the second step. So, nobody can acquire justification for Table-3 or Election-3 via these arguments - the threat of "easy justification" disappears. Thus an advantage of the conservative diagnosis of Moorean anti-skepticism is that it meshes well with our judgments about everyday cases.

This conservative account of what is wrong with Moorean reasoning is extremely appealing. In this paper, however, I intend to defend Moorean reasoning. Does this mean I will argue that conservatism is false? No, not exactly. I will argue that conservatism is *ambiguous* because it makes a claim about "justification to believe" and different philosophers mean different things when they use this term. I have no intention of defending any particular interpretation of "justification to believe" here. Instead, I will argue that there is no interpretation of "justification to believe" wherein conservatism is both true and inconsistent with Moorean reasoning being good reasoning. Either way, Mooreans are out of the woods: conservatism is either false or harmless.

The structure of the paper is as follows: part 1 will discuss what is at issue in the debate over Moorean reasoning. Part 2 will lay out the terms of the debate more carefully, and part 3 contains a discussion of "justification to believe" that will allow me to distinguish two possible

---

[4] Cohen and Wright (2008) offer versions of this worry about Moorean reasoning; Cohen discusses easy knowledge, rather than easy justification, but the basic shape of the worry is the same.

version of conservatism. The next two sections will offer a dilemma: part 4 will argue that one of the two versions of conservatism is implausibly strong, and part 5 will show that the other version cannot explain what is wrong with Moorean reasoning. The remainder of the paper will tie up loose ends: part 6 will discuss a possible objection to my argument and clarify the scope of cases in which Moorean reasoning is appropriate, part 7 will use the material developed in parts 3 through 5 to show what is wrong with a well-known argument against Mooreanism in formal epistemology, and part 8 will conclude by discussing rational agnosticism in light of distinctions made in part 3.

2.1   WHY DOES IT MATTER WHETHER MOOREAN REASONING IS USELESS?

Why does it matter what we say about Moorean reasoning? First, and most famously, what we say about Mooreanism constrains what we say about evil demon-style skeptical arguments. Jim Pryor has argued that a very common way of presenting such arguments is not at all compelling.[9] Consider:

> **Premise:** (Simple-1) I do not have justification to believe that I am not being deceived by an evil demon
> **Premise:** (Simple-2) If I do not have justification to believe that I am not being deceived by an evil demon then I do not have justification to believe that I have hands
> **Therefore:** (Simple-3) I do not have justification to believe that I have hands

It is not clear that anyone should lose any sleep over this argument. The first premise is itself a skeptical conclusion, and hardly something that most people accept pre-philosophically. So it

---

[9] This discussions of the shortcomings of the simple skeptical argument, and the existence of a more powerful alternative, follows Pryor (2000).

looks like we are free to reject it. Yet without Simple-1, skeptics cannot make any mischief with this style of argument.

However, skeptics can produce a more powerful argument that does not rely on anything so contentious as Simple-1. The really worrying skeptical argument is:

> **Premise:** (Nasty-1) Either I do not have justification to believe that that I am being deceived by an evil demon *or* I do have justification to believe it, but my justification is based (at least in part) on beliefs justified by perception
> **Premise:** (Nasty-2) If I have justification to believe any proposition on the basis of perception, then I must have independent (of perception) justification to believe that I am not being deceived by an evil demon
> **So:** (Nasty-3) I do not have justification to believe that I am not being deceived by an evil demon
> **Premise:** (Nasty-4) If I do not have justification to believe that I am not being deceived by an evil demon then I do not have justification to believe that I have hands
>
> **Therefore:** (Nasty-5) I do not have justification to believe that I have hands

The third step of this argument follows from the first two because Nasty-1 says that any justification I have to believe that I am not being deceived by an evil demon must depend upon perception, while Nasty-2 says that justification to believe *anything* on the basis of perception requires perception-independent justification to believe that I am not being deceived by an evil demon. More simply, Nasty-1 says that justification to believe that I am not being deceived could *only* come from perception, while Nasty-2 says that justification to believe that I am not being deceived *cannot* come from perception. Nasty-3 follows because there is nowhere left for justification for the anti-skeptical claim to come from.

Nasty-1 is much weaker than Simple-1. More importantly, Nasty-1, unlike Simple-1, is not a statement that most anti-skeptics can flatly reject. So this argument looks sound—or in any case, the range of options for responding to skepticism is narrow—if Nasty-2 is true. But Nasty-2 is just the claim that Moorean reasoning, at least when it comes to perception, is no good.

The situation regarding inductive skepticism is analogous. Consider a familiar sort of argument:

> **Premise:** (Simple Induction-1) I do not have justification to believe that induction will be reliable in the future
> **Premise:** (Simple Induction-2) If I do not have justification to believe that induction will be reliable in the future then I do not have justification to believe that the sun will rise tomorrow

Therefore: (Simple Induction-3) I do not have justification to believe that the sun will rise tomorrow

Non-skeptics can comfortably deny Simple Induction-1, so this argument is not very powerful. The more worrying argument is:

> **Premise:** (Nasty Induction-1) Either I do not have justification to believe that that induction will be reliable in the future *or* I do have justification to believe it, but my justification is based (at least in part) on beliefs justified by induction
> **Premise:** (Nasty Induction-2) If I have justification to believe any proposition about the future on the basis of induction, then I must have independent (of induction) justification to believe that induction will be reliable in the future
> **So:** (Nasty Induction-3) I do not have justification to believe that induction will be reliable in the future
> **Premise:** (Nasty Induction-4) If I do not have justification to believe that induction will be reliable in the future then I do not have justification to believe that the sun will rise tomorrow.
> **Therefore:** (Nasty-5) I do not have justification to believe that the sun will rise tomorrow

There is no premise here that non-skeptics can flatly reject. However, Mooreans reject Nasty Induction-2, since they think I can acquire justification to believe that induction will be reliable in the future via inductive inference. Once again, Mooreanism is the most promising way to deflect a very powerful skeptical argument.

The rejection of Moorean reasoning thus plays a crucial—if often tacit—role in some perennially vexing skeptical problems. If we are going to get a handle on responding to these skeptical arguments, we must figure out whether Moorean reasoning is epistemically useless. Of course, philosophers who reject Mooreanism need not be skeptics. Non-Mooreans can avoid skepticism by embracing **rationalism** and saying that we can acquire justified beliefs in a host of important anti-skeptical claims—including *I am not a handless brain in a vat* and *induction will be reliable in the future*—via apriori, armchair reasoning. The most compelling kind of rationalism says that apriori justification to believe in anti-skeptical claims does not rely on the use of *any* belief-forming method; rationalists can thus deny Nasty-1 and Nasty Induction-1 by saying that each of us has default justification to believe that she is not in a skeptical scenario.[10]

---

[10] The conservative anti-Mooreans White and Wright both embrace some version of rationalism. Wright uses somewhat unconventional nomenclature, however: he refers to apriori "entitlement" to believe anti-skeptical claims; he reserves the term "justification" for bits of entitlement that we somehow earn, rather than possess by default.

Discussion of the merits and drawbacks of rationalism is a large topic, well beyond the scope of this paper; my aim here is to show how to defend Moorean reasoning and thereby make our menu of anti-skeptical options seem more appealing.

One way of viewing the space of options here is that everyone must bite one bullet: we must choose between skepticism, rationalism, and Mooreanism. If that is correct, there might be a "best fit" argument for Mooreanism. If we can make rationalism and skepticism look bad enough, holding our nose and denying conservatism—and thus embracing Mooreanism—can look like the best option. Fortunately, this bleak picture is incorrect: a more careful examination of conservatism will reveal that it gives us no reason to reject Mooreanism.

2.2 WHAT, EXACTLY, IS CONSERVATISM?

We need to make three concepts more precise before stating conservatism carefully. The first concept is **undermining**. Suppose Samantha has some evidence, E, for a proposition p; E is strong enough evidence to lend Samantha justification to believe that p.[11] I'll use the shorthand Support(E,p) to mean the rational support that E lends to p. A **defeater** for p is a proposition such that, when Samantha comes to reasonably believe that it is true, she no longer has justification to believe that p. One type of defeater is an **outweighing defeater**: a proposition that provides evidence for the falsity of p and thus renders belief in p inappropriate in spite of the reasons given by E for counting p true. An outweighing defeater leaves Support(E,p) unchanged. Another type of defeater is an **underminer**. Learning that an underminer is true eliminates Support(E,p). For instance, suppose I see a happy looking crowd emerging from Shea stadium one summer night. That evidence gives me justification to believe that the Mets just won a game. Suppose I read in the paper the next day that the Mets lost: my justification to

---

Cohen eschews both rationalism and Mooreanism in favor of holism; discussion of Cohen's positive view—which makes a distinction between two levels of knowledge—would take us very far afield.

[11] I won't assume that all evidence is propositional; as I'll use the term, sense experience can be evidence.

believe that they won has been *outweighed*. Suppose instead that I ask someone in the crowd and they tell me that they came from a rock concert and the Mets played an away game: my justification to believe that the Mets just won has been *undermined*.[12] I'll call a proposition an **anti-underminer** if its negation is an underminer.

The second concept is **justification to believe**, which I'll also call **propositional justification**. A statement such as "p is justified for Stuart" is ambiguous. One sense of that phrase is: Stuart has adequate evidence and/or whatever else is needed to provide rational support for the proposition that p. That is *propositional justification*, or justification to believe that p. This definition is a bit imprecise and we will consider some ways of making it more exact later on, but it will suffice for now. A second sense is: Stuart has an appropriate belief that p. This means that, first, Stuart believes that p; second, Stuart has propositional justification for p; third, Stuart's belief that p is **based**—that is, psychologically dependent upon—the evidence (or whatever else) that lends him justification to believe that p. When these three conditions are met, Stuart's belief that p is **doxastically justified**.

The third concept is **independent** propositional justification. Though it is easy to gesture at the rough sense in which *independent justification* is used in formulating conservatism, it is very difficult to define independence precisely. I will not attempt a precise definition here. As a rough definition—good enough for our purposes—we can say that someone has p-independent justification to believe a proposition if it is possible for someone with his evidence to reason his way to a justified belief in that proposition without going through p as an intermediate step.

We are now in a position to give a careful statement of conservatism:

> **Conservatism:** S can acquire a justified belief that p on the basis of E only if S has p-independent justification to believe each of the anti-underminerS for Support(E,p).[13] [14]

---

[12] Many propositions diminish Support(E,p) to some degree. I will use the term *underminer*, though, only to refer to propositions that eliminate Support(E,p) entirely.

[13] The formulation I have given is in terms of all-or-nothing justification. However, fans of degrees of justification will want to endorse a slightly more general formulation of conservatism, namely:

> **Conservatism***: S can additional rational support for p on the basis of E only to the degree that he has p-independent justification to believe anti-underminers for Support(E,p).

For the sake of simplicity, I will stick to the all-or-nothing formulation; the difference won't matter for the purposes of this paper.

[14] Not everyone who writes about this issue defines liberalism and conservatism as I do here: while I have represented conservatism as a conditional,

I will call anyone who denies conservatism a **liberal**.[15] It is important not to overstate the difference between liberalism and conservatism. The two positions are closer than they may initially appear in two ways. First, liberals may agree that S can only acquire additional support to believe that p on the basis of E if she has p-independent justification to believe *some* of the relevant anti-underminers, or that she *sometimes* needs justification for all of the relevant anti-underminers. Second, both liberals and conservatives agree that when someone ought to believe that an underminer for Support(E,p) is true, he cannot acquire justification to believe that p on the basis of E. All parties also agree that when someone ought to believe that all of the underminers for Support(E,p) are false there is no obstacle for him acquiring justification for p on the basis of E. Liberalism and conservatism give different verdicts *only* in cases where someone is rationally agnostic about at least one relevant underminer. The disagreement between liberals and conservatives centers on the force of rational agnosticism: is being rationally agnostic about an underminer for Support(E,p) enough to threaten the support that E lends to p?

I will call a bit of reasoning **Moorean** if it involves acquiring a belief that p on the basis of E, and then involves acquiring a belief in some anti-underminer for Support(E,p) on the basis of p; I will call someone a Moorean if he thinks that Moorean reasoning is a way to acquire justification to believe its conclusion for the very first time.

(S can acquire support to believe p on the basis of E)$\rightarrow$(S has p-independent justification to believe all of the anti-underminers for Support(E,p))

it is common to represent conservatism as some sort of claim involving "because" or "in virtue of", such as:

S can acquire support to believe that p on the basis of E only *in virtue of* having p-independent justification to believe all of the anti-underminers for Support(E,p)

I will focus on the conditional since it is unclear to me just what "in virtue of" means in this context. My way of formulating the issue simplifies the logical space. For instance, Nico Silins (2008) describes his view as non-Moorean liberalism. He argues against conservatism by saying that rational support to believe that p need not be *in virtue of* justification to believe the relevant anti-underminers, yet he does embrace conservatism in my sense. On my way of describing thing, Silins is just a conservative. Indeed I think Silins's argument shows why, obscurity aside, the *in virtue of* claim does not carve at the relevant joints: since the conditional is adequate to block Moorean arguments, making the *in virtue of* claim needlessly exposes conservatism to additional objections.

[15] This terminology follows Pryor (2004) and (2008). It is common to use the term "dogmatism" to refer to what I call liberalism; strictly speaking, dogmatism refers to a species of anti-conservatism about perceptual beliefs only.

To get clear on just what conservatism means, we need to pin down the notion of propositional justification more precisely; once we have made a few more distinctions, we will be in a position to pose a dilemma involving the meaning of "propositional justification" in the formulation of conservatism.

2.3   JUSTIFICATION AND TWO KINDS OF EVIDENCE

Before nailing down the different senses of propositional justification, I need to make a pair of observations about three other notions: first, appropriate—or doxastically justified—belief; second, evidence; third, good reasoning.

The first observation is about the relationship between evidence and doxastically justified belief. There are two principles about the epistemic role of evidence that are widely considered platitudes: first, that people ought to apportion their degrees of confidence according to the evidence, and second that some fact is evidence for (or against) p only if it "bears on" the truth of p or indicates that p is more (or less) likely to be true. It turns out, though, that these two principles sit very uneasily with one another. Often someone ought to adjust her confidence in p upon learning facts that do not indicate that p is any more or less likely to be true. Consider:

The Unsuccessful Detective: Detective Smith is investigating a murder; his current theory is that the butler did it. Detective Smith has a bunch of evidence that the butler did it: the butler had a motive, his alibi was uncorroborated, footprints in the victim's blood match the butler's shoes, and so on. On the basis of this evidence, Detective Smith becomes very confident that the butler did it. Then, at a slightly later time, Detective Smith learns something: he has an incredibly poor track record of solving cases. A new, extremely reliable DNA testing technology reveals that his theories in a huge number of previous cases—cases in which he was quite confident that he knew who had committed the crime—turn out to have been wrong. When Detective Smith learns that he has a poor track record, he becomes a good deal less confident that the butler did it.

It is clearly reasonable for Detective Smith to revise his confidence down in light of the new evidence about his track record – indeed, it would be unreasonable for him *not* to do so. But this

thought can appear puzzling: after all, we might think that a hallmark of epistemic rationality is that one should revise one's confidence in a proposition p only on the basis of acquiring evidence that is relevant to whether p. Yet, the new evidence—that Detective Smith has a poor track record—is irrelevant to whether the butler did it. Certainly if I were investigating the murder, I would want to know what the crime scene looked like, what the butler's alibi was, and so on: but whether Detective Smith is good at his job wouldn't be relevant information. What's going on here?

The evidence that determines how confident Detective Smith ought to be that the butler did it falls into two categories. The first category contains evidence that is relevant to whether the butler did it: the butler's alibi, motive, footprints, and so on. Call this **object-directed** evidence, since it bears on the object of Detective Smith's attitude. A second type of evidence does not bear at all on the object of Detective Smith's attitude, yet does affect how confidently he can hold that attitude: in this case, the revelation that Detective Smith has a poor track record. Call this **reasoning-directed** evidence, since it affects the rational confidence that Detective Smith can have in p not by bearing on whether p, but rather by bearing on whether he is able to figure out whether p via reasoning from the clues of the case. Detective Smith ought to adjust his attitude towards whether the butler did it in response to both kinds of evidence. The first observation, then, is that rationally responding to one's evidence means taking both object- and reasoning-directed evidence into account.[16]

The second observation concerns the relationship between evidence and routes of good reasoning. Reasoning-directed evidence does not, typically, just diminish the rational confidence that someone ought to have in a proposition *tout court*; rather, it diminishes the confidence that someone ought to have in a proposition given that it rests on a certain kind of reasoning. For instance, evidence of Detective Smith's poor track record in determining the culprit via examining the clues should lower his confidence that the butler did it insofar as that conclusion is based on his reasoning from the clues. However, if a reliable source with a good track record—

---

[16] David Christensen, in his unpublished manuscript "Higher-Order Evidence", describes the distinction between what I call object-directed and reasoning-directed evidence at some length. He uses the term "higher-order evidence" to refer to what I call reasoning-directed evidence. I think his nomenclature is seriously misleading: "higher-order" evidence sounds like evidence about evidence, whereas what he is talking about is evidence about one's reasoning ability. Beyond that quibble, though, his discussion of the rational response to evidence of fallibility is excellent and far more thorough than my quick remarks here.

Sherlock Holmes, say—tells Detective Smith that the butler did it, Smith can be quite confident that the butler did it. His reasoning-directed evidence gives no cause to doubt his ability to form true beliefs via the testimony of others. Reasoning-directed evidence can defeat reasoning from one body of evidence—say, the clues—to a proposition but leave reasoning from another body of evidence—say, Holmes's testimony—untouched. The second observation, then, is that reasoning-directed evidence can generate "fine grained" defeat: that is, it can defeat any path of reasoning from one body of evidence to a given proposition while leaving paths of reasoning from other bodies of evidence to that very same proposition unscathed.

In sum, there are three kinds of evidence that can make someone rationally less confident that p:

1. **Outweighing defeaters:** object-directed evidence that ~p.

2. **Undermining defeaters:** object-directed evidence that some evidence E does not indicate the truth of p. Underminers for Support(E,p) should lower someone's confidence in p insofar as that confidence is based on E – someone who believes that p for another reason shouldn't alter her confidence in light of acquiring undermining evidence for Support(E,p).

3. **Reasoning-directed defeaters:** evidence that one shouldn't trust one's ability to figure out whether p on the basis of E. This, too, should only lower someone's confidence in p insofar as it is based on E. But it only has this effect on one person – other people who believe that p on the basis of E shouldn't lower their confidence one bit when they discover that Detective Smith's ability to figure out whether p on the basis of E is not so hot.[17]

Note that the same fact can be more than one type of evidence, depending on who learns it and what else they believe. For instance, the fact that Detective Smith has a poor track record is reasoning-directed evidence for Detective Smith. However, if I believe that the butler did it based on the testimony of Detective Smith, then for me this evidence is an underminer—a kind

---

[17] We can differentiate the three types of defeater in formal terms. Call O an outweighing defeater for p, U an undermining defeater for Support(E,p), and R a reasoning-directed defeater for S concluding that p on the basis of E. $Pr(p|O) < Pr(p)$, whereas $Pr(p|U$ and $E) = Pr(p)$. Reasoning directed defeaters behave quite differently: $Pr(p|E$ and $R) = Pr(p|E)$, yet my confidence in p upon learning E and R should be lower than my confidence in p upon learning E alone. See section 7 for further discussion of this point.

of object-directed evidence—for the support that Detective Smith's verdict gives for the conclusion that the butler did it.

With our two observations in mind we can ask about the relationship between propositional justification, evidence, and good reasoning. In particular, does having propositional justification for p always entail the existence of a sound, undefeated reasoning route from one's evidence to the conclusion that p? I will call views of propositional justification wherein the answer is " yes" **robust views of propositional justification** and views wherein the answer is "no" **anemic views of propositional justification**. The two types of view differ in cases where S has good object-directed evidence for a proposition p, but there does not exist a sound, undefeated deliberative route from his evidence to p since all such routes are blocked by reasoning-directed defeat (they may differ in other cases as well, but I'll focus on examples involving reasoning-directed defeat). In cases like this, anemic views can say that S has propositional justification for p, but robust views must say that he lacks propositional justification for p.

We can plug each type of view of propositional justification into conservatism and get either:

Robust Conservatism: S can acquire a justified belief that p on the basis of E only if there is some p-independent sound, undefeated bit of reasoning from S's evidence which concludes in doxastically justified beliefs in each of the anti-underminers for Support(E,p).

or

Anemic Conservatism: S can acquire a justified belief that p on the basis of E only if he has p-independent justification to believe each of the anti-underminers for Support(E,p), where this justification need not involve the existence of a sound, undefeated bit of reasoning from S's evidence to a doxastically justified beliefs in the anti-underminers.

These two versions of conservatism agree that someone can acquire justification to believe that p on the basis of E only if he has p-independent epistemic grip on relevant anti-underminers. What they disagree about is how strong this epistemic grip needs to be. Robust Conservatism says the grip must be fairly strong. There must be a good piece of p-independent reasoning from one's evidence to the relevant anti-underminers; moreover, one's evidence must *not* include anything that provides reasoning-directed

defeat for beliefs in the relevant anti-underminers. Anemic Conservatism denies that one's grip on the relevant anti-underminers need be quite this strong; in particular, Anemic Conservatives are happy to countenance acquiring a justified belief that p on the basis of E while having reasoning-directed defeaters that prevent forming a justified belief in some anti-underminer for Support(E,p). This minor difference in how the two views treat reasoning-directed defeat has surprisingly strong ramifications for what each type of conservative ought to say about Moorean reasoning.

I do not think there is any answer to the question "Which of these is the *correct* conception of propositional justification?" The term "propositional justification" is just a bit of jargon, and philosophers can use it however they want. I suspect that most epistemologists have something like Robust Conservatism in mind when they use the term, but I doubt whether there is universal agreement. In the next two sections I will argue that neither of these two versions of conservatism can show what is wrong with Moorean reasoning: Robust Conservatism is false, and Anemic Conservatism is consistent with Moorean reasoning being good reasoning. Since the robust and anemic views exhaust the logical space for understandings of propositional justification— propositional justification for p either entails the existence of a good bit of reasoning that concludes that p or it doesn't—this creates a dilemma for fans of the view that conservatism can tell us what is wrong with Moorean reasoning.

2.4   AGAINST ROBUST CONSERVATISM

Robust Conservatism says that if S can acquire justification to believe that p on the basis of E, then there must be sound deliberative routes from his evidence to each of the anti-underminers for Support(E,p). Any counterexample to Robust Conservatism must have the following two features:

1) There is no sound, undefeated bit of reasoning from S's evidence that concludes in a doxastically justified belief in some anti-underminer for Support(E,p).

2) S is clearly able to acquire a justified belief that p on the basis of reasoning from E.

It is easy to formulate such counterexamples. Consider:

The Mediocre Mathematician: Stan knows himself to be mediocre at math. One day a friend asks him if, in general, assuming that a proposition is false and then deriving a contradiction is sufficient to establish that the proposition is true. Stan, quite appropriately, says he's stumped: he just has no idea how to assess if a method of proof is kosher. He is rationally agnostic about whether *reductio ad absurdum* is a valid method of argument. Later Stan is wondering whether there are any even prime numbers besides 2. He thinks to himself: *suppose n is an even prime number (and not 2), then, since n is even, it is divisible by 2, but since it is prime (and not 2), it is not divisible by 2.* He concludes that there are no even primes besides 2 – and this belief is justified.

The Persuasive Lecture: Seth becomes convinced by a persuasive lecture by a respected authority on ethics who claims that while we should trust our ordinary judgments about the rightness or wrongness of individual acts, it is impossible to tell whether our faculties for making moral judgments are reliable (since the only yardstick we can use to measure their accuracy is our own judgments). Seth can't really follow the details of the arguments, but since the lecturer is respected, he concludes that it must be true that we can trust our ordinary moral judgments and that nobody can say whether his own faculties for making moral judgments are reliable. Seth sees some children setting a cat on fire, and forms a justified belief that they are acting wrongly, while being rationally agnostic about the proposition *my faculties for making moral judgments are reliable*.

The Seasoned Interviewer: Sasha has interviewed many criminal suspects for the police, and she is very good at her job. After much practice and study, she is excellent at telling truth from lies and she knows it. One day a colleague asks Sasha if a suspect who frequently touches his nose during an interview is more likely to be lying. Sasha replies, truthfully, that she isn't sure; she finds it very hard to articulate such rules in the abstract since her previous attempts to do so have been unsuccessful. Shortly thereafter, Sasha interviews a suspect who frequently touches his nose. Sasha acquires a justified belief that the suspect is lying, and her belief is based, in large part, on his frequent nose-touching.

This list could go on. While any one example might be open to objection, the pattern is clear. In each of these cases, someone is able to form a justified belief that p on the basis of E while being rationally agnostic about some anti-underminer for Support(E,p). Robust Conservatism is false. All of these counterexamples have the same structure: there is a reasoning-directed defeater that

prevents any reasoning from S's evidence to some anti-underminer from being good reasoning. It is reasoning-directed evidence that blocks any good reasoning from Sasha's evidence to *frequent nose-touching is a sign of dishonesty*, from Seth's evidence to *my faculties for making moral judgments are reliable*, and from Stan's evidence to *reductio ad absurdum is a valid argument form*. Yet these reasoning-directed defeaters do not block paths of good reasoning concluding *the suspect is lying*, *setting a cat on fire is wrong*, or *there are no even prime numbers greater than 2*.

The problem for Robust Conservatism is that reasoning-directed evidence can be fine-grained: it can thus defeat good reasoning concluding in one proposition while leaving reasoning towards some nearby propositions untouched. For instance, in the case of the Mediocre Mathematician, reasoning-directed evidence defeats any reasoning from Stan's evidence to complicated or difficult-to-establish mathematical claims, but not to simple or easily grasped ones. He thus reasonably doubts his ability to figure out whether *reductio ad absurdum* is kosher, but does not doubt his ability to conduct simple proofs of claims such as *there are no even primes greater than 2*.[18] The point is not that it's harder to have propositional justification for complex or difficult-to-establish mathematical claims. Rather, the point is that *if you have evidence* that you cannot reliably figure out whether complex or difficult-to-establish mathematical claims are true—as most of us do, and Stan does in spades—*then* it's harder to have propositional justification for complex or difficult-to-establish mathematical truths.

It's worth distinguishing the problem I have identified here from a common—but unsuccessful—complaint about conservatism. Some philosophers worry that conservatism makes it impossible for all but the most cognitively sophisticated thinkers to acquire rational support to believe much of anything. Let's make this worry concrete with an example:

The Naïve Reasoner: Sarah is a normal child of age 4. One day Sarah sees that it is raining outside and she forms the belief that the ground is wet. Sarah does not have any attitude about the underminer *there is a giant tarp covering the ground outside*. Indeed, she lacks the concept TARP and so has no beliefs about the presence or absence of tarps.

---

[18] One might grant this point, but still think that agnosticism about *reductio ad absurdum* ought to somewhat diminish one's confidence in *there are no even primes greater than 2*. I do not disagree with this thought. All that matters is that belief in *there are no even primes greater than 2* is doxastically justified in spite of rational agnosticism about the anti-underminer; the counterexample does not require that this justification is completely unscathed by rational agnosticism about the anti-underminer.

Here's why this case might be trouble. Sarah's belief that the ground is wet is justified. The proposition *there is no giant tarp covering the ground outside* is an anti-underminer for Sarah acquiring justification to believe that the ground is wet based on the evidence that it is raining. So, conservatism says that Sarah must have independent (of her belief that the ground is wet) justification to believe that *there is no giant tarp covering the ground outside* if she can acquire justification to believe that the ground is wet given her evidence that it is raining. Yet Sarah cannot so much as entertain that anti-underminer: it involves the concept TARP, which she lacks.

This example does not cause any problems because conservatives need not claim that Sarah needs to *believe* all of the relevant anti-underminers, or even that she be capable of believing them. All conservatives need to say is that Sarah must have propositional justification for anti-underminers, which she has. There is a sound, undefeated deliberative route from Sarah's evidence—which includes the fact that the ground is usually uncovered—to the conclusion that *there is no giant tarp covering the ground outside* via inductive inference. Sarah cannot perform that inference, but it is no part of conservatism to insist that she be able to do so. If conservatives said that Sarah needed doxastically justified beliefs in the relevant underminers, this case would present trouble; but they don't, so it doesn't.

What is the difference between this unsuccessful objection and my own objection? The unsuccessful objection accuses conservatism of having implausible implications in cases where someone's impoverished capacities prevent her from being able to believe an anti-underminer. My objection accuses conservatism of having implausible implications in cases where someone's *evidence* is inconsistent with appropriately believing an anti-underminer. In the case where impoverished capacities prevent the formation of a doxastically justified belief, Robust Conservatives can say that Sarah has propositional justification for *the ground is wet*. In the earlier cases—where evidence defeats all of the available routes of reasoning to the relevant anti underminers—they cannot.

2.5   ANEMIC CONSERVATISM TO THE RESCUE?

The counterexamples in the previous section all turned on plugging the robust conception of propositional justification into conservatism; if we replace that view of propositional justification with a weaker one, we get some version of:

Anemic Conservatism: S can acquire additional rational support to believe that p on the basis of E only if he has p-independent justification to believe each of the anti-underminers for Support(E,p), where this justification need not involve the existence of a sound, undefeated bit of reasoning from S's evidence to a doxastically justified belief that p.

Of course, Anemic Conservatism is not so much a particular view but a family of views. To get a positive story about just what conservatism means, we would also need some necessary or sufficient conditions for having propositional justification; as it stands, all Anemic Conservatism tells us about propositional justification is what it does *not* entail. Many forms of Anemic Conservatism can survive the counterexamples of the previous section. However, all forms of Anemic Conservatism are consistent with Moorean reasoning sometimes being good reasoning. The problem cases arise when someone is in a position where:

> (a) He has p-independent justification to believe that ~U in the anemic sense
>
> (b) There is no p-independent sound reasoning from his evidence which concludes in a doxastically justified belief that ~U; that is, he lacks p-independent justification to believe that ~U in the robust sense.
>
> (c) He can form a doxastically justified belief that ~U via reasoning from p, perhaps along with other premises. That is, he can have p-dependent justification to believe ~U in the robust sense.

If someone is in this position, prior to acquiring E there is no way for anyone with his evidence to form a doxastically justified belief that ~U (that's condition (b)). However, when he acquires evidence E, he can form a justified belief that p (that follows from (a) and Anemic

Conservatism). He can then form a doxastically justified belief that ~U via inference from p (that's condition (c)). But that's just Moorean reasoning.

As a concrete example, consider a particular species of Anemic Conservatism:

Object-Directed Conservatism: S can acquire additional rational support to believe that p on the basis of E only if he has, independent of p, adequate object-directed evidence for each of the anti-underminers for Support(E,p).

By *adequate object-directed evidence* for an anti-underminer I mean enough object-directed evidence for someone—perhaps someone with different conceptual capacities or different reasoning-directed evidence—to form a doxastically justified belief in that anti-underminer. Here is a case where Object-Directed Conservatism licenses Moorean reasoning:

The Seasoned Interviewer Redux: As before, Sasha is an excellent and experienced police interviewer. Also as before, Sasha sincerely pleads agnosticism when a colleague asks her if a suspect who frequently touches his nose during an interview is more likely to be lying – she finds it hard to reason about such rules in the abstract and has a poor track record when doing so. Shortly thereafter, Sasha interviews a suspect who frequently touches his nose. Sasha forms a justified belief that the suspect his lying and this belief is based, in large part, on his frequent nose-touching. After the interview, Sasha remembers her colleague's question. She realizes that nose-touching seemed to her a clear sign that the subject was lying. She concludes that frequent nose-touching is a sign of dishonesty after all.

Sasha has engaged in some Moorean reasoning here. According to Object-Directed Conservatism, Sasha's reasoning is kosher. Here is why:

(a') Prior to conducting her interview with the nose-touching suspect, Sasha has adequate object-directed evidence to believe that *frequent nose-touching is a sign of dishonesty* (and, we will assume, all of the other anti-underminers for Support(the suspect is touching his nose frequently, the suspect is lying)) so according to Object-Directed Conservatism there is no obstacle to Sasha acquiring a justified belief that the suspect is lying based on his frequent nose-touching.

(b') Prior to conducting her interview, there is no sound, undefeated bit of reasoning from Sasha's evidence to the conclusion *frequent nose-touching is a sign of dishonesty* – all such routes are defeated by Sasha's evidence that she is no good at reasoning about such rules in the abstract.

(c') Sasha can form a doxastically justified belief in *frequent nose-touching is a sign of dishonesty* via reasoning from *the suspect is lying* and the knowledge that her belief that the suspect is lying is based upon his frequent nose-touching.

So, prior to conducting the interview, Sasha is rationally agnostic about *frequent nose-touching is a sign of dishonesty*. But, once she does conduct the interview, she can form a justified belief that the suspect is lying, and then a justified belief in *frequent nose-touching is a sign of dishonesty*. There is nothing wrong with this Moorean pattern of reasoning – at least, not according to Object-Directed Conservatism.

Nothing I have said here should be construed as an attack on Anemic Conservatism. I strongly suspect that some form of Anemic Conservatism is true. Rather, my point is that the incompatibility of conservatism and Mooreanism is not so easy to establish. The incompatibility depends upon plugging the robust conception of propositional justification into conservatism. Anemic Conservatism allows someone to form a belief in p on the basis of E without the existence of a sound, undefeated bit of reasoning from his evidence to some anti-underminers for Support(E,p). However, once he has justification to believe that p, new, undefeated routes of reasoning to beliefs in anti-underminers open up. It's precisely the anemia that licenses Moorean reasoning.

2.6   SKEPTICISM AND MOOREANISM

So far I have shown that Robust Conservatism delivers false results in situations involving reasoning-directed defeat and that Anemic Conservatism is too weak to block Moorean reasoning. So there is no version of conservatism that is both true and capable of explaining what is wrong with Moorean reasoning. But perhaps foes of Mooreanism can pull off a strategic retreat with something like the following line:

> Robust Conservatism is indeed false; it badly misfires in cases involving reasoning-directed defeat. But that gives us no reason to believe that Robust Conservatism is not at least roughly correct – we should not abandon it it, but rather insert a proviso that it does not apply in cases where reasoning-directed evidence blocks justification to believe an anti-underminer.

I do not think this line of retreat is worryingly *ad hoc*: the only cases where robust and anemic versions of conservatism diverge involve reasoning-directed defeat, so it is natural to wonder whether we can set such cases aside and keep an otherwise promising position more-or-less intact. Everything that I have said so far is consistent with conservatism being correct in cases that do not involve reasoning-directed defeat; indeed, I strongly suspect that it is correct in such cases. However, while this retreat might help to make conservatism incompatible with some Moorean reasoning, it will do no good in the important cases. Call the cases where this line of retreat admits that Robust Conservatism misfires—cases where reasoning-directed evidence blocks justification to believe an anti-underminer—*problem cases*. Fans of Mooreanism as an anti-skeptical strategy can point out that the skeptical arguments discussed in part 2 are problem cases. Recall:

> **Enabling condition: (**Moore-1) [an experience as of a hand]
> **So:** (Moore-2) I have a hand
> _____
> **Therefore: (**Moore-3) I am not a handless brain in a vat

Earlier we saw two prominent non-skeptical diagnoses of this argument: Mooreans think that this reasoning can generate justification to believe the conclusion *for the first time* whereas conservative rationalists think that someone can perform this reasoning only if he *already* has justification to believe the conclusion. But we should now think to ask rationalists whether this is justification in a robust or anemic sense. If it's anemic, Moore's proof can still be a way to acquire a justified belief in its conclusion for the first time. If it is robust, there must have been a prior, p-independent route of sound, undefeated reasoning from Moore's evidence to p. If rationalism is to be an alternative to Mooreanism, alternative Moorean routes of reasoning—routes that involve the use of beliefs acquired via perception—won't do. So the alternative route must consist of armchair reasoning: the sort of reasoning one can engage in without any perceptual inputs. But I doubt very much whether any such route is available. In particular, such a route could only provide Moore with *robust* justification to believe that he is not a handless brain in a vat if his reasoning-directed evidence is consistent with trusting himself to reliably determine which deeply contingent propositions about his surroundings are true based on pure armchair reflection. But I don't think any of us has reasoning-directed evidence consistent with trusting ourselves to figure out what the actual world happens to be like by armchair reflection

alone. So this is a problem case: there is reasoning-directed defeat that prevents anyone from determining that he is not a handless brain in a vat from the armchair. If Robust Conservatism misfires when dealing with problem cases, Robust Conservatism misfires when dealing with Moore's proof itself.

A similar story applies to Induction 1-3. Perhaps we have evidence about what will happen in the future. But none of us, I think, has reasoning-directed evidence consistent with trusting herself to figure out what the future will be like *without using induction*. So this case, too, is a problem case: even if we have justification to believe that induction will be reliable in the future in some anemic sense, our reasoning-directed evidence is inconsistent with having such justification in a robust sense unless we can acquire it using Moorean reasoning.

The upshot is that the retreat can, at best, establish that Moorean reasoning from E to p and then from p to ~U is epistemically useful when, and only when, reasoning-directed defeaters block alternative reasoning routes to ~U but do not block reasoning from p to ~U. In day-to-day life, these situations are rare. Most of the time Moorean reasoning is indeed useless. Recall our earlier pedestrian examples of Moorean reasoning:

> **Premise:** (Table-1) This table looks red
> **So:** (Table-2) This table is red
> **Therefore:** (Table-3) This is not a white table under red lights

> **Premise:** (Election-1) Someone just checked a box on a piece of paper and slid it into a slot
> **So:** (Election-2) Someone just voted
> **Therefore:** (Election-3) An election is happening, rather than a rehearsal

The diagnosis on offer says that these bits of reasoning are indeed epistemically useless, since none of us has reasoning-directed evidence that she cannot determine whether the lighting in a room is red or whether an election is occurring by any non-Moorean methods. So, there is no threat of "easy justification": Moorean reasoning really is epistemically useless in these everyday cases. However, in rare cases Moorean reasoning is fine; as it happens, these rare cases include responses to global skeptical hypotheses. Moorean reasoning is indeed usually bad reasoning: philosophers who look askance at Moore's proof are right about that. However, Moorean reasoning is fine in narrowly-defined cases – and, in particular, it is fine when skeptics generate reasoning-directed doubt about anti-underminers.

So much, then, for showing that neither version of conservatism can block Moorean reasoning. Fans of Mooreanism may worry about the strategy of just conceding that some form of Anemic Conservatism is true on the grounds that this strategy can avoid skepticism only by conceding that some form of rationalism is true. After all, Moorean reasoning is only consistent with Anemic Conservatism for those who have antecedent anemic justification to believe *I am not a handless brain in a vat*. Such justification must be antecedent to all perceptual beliefs, i.e. apriori. Part of the appeal of Mooreanism is that allows a non-rationalist solution to skeptical worries; my solution just seems to retreat on this front.

Of course, nothing I have said here requires that Anemic Conservatism is true; I don't have a good argument against it, but everything I have said here is consistent with its falsity. In addition, it is not clear how bad anemic apriori justification to believe anti-skeptical hypotheses is: granting that such anemic rationalism is true may not be a large concession since it is consistent with saying that *nobody* can form justified beliefs in anti-skeptical hypotheses without evidence acquired from experience. It is not at all clear to me that, whatever the best arguments against rationalism are, they make trouble even for anemic forms of rationalism. Finally, though, there are some considerations in formal epistemology that strongly suggest that some form of anemic rationalism must be true; I will look at these arguments in some detail in the next section.

2.7   BAYESIANISM AND CONSERVATISM

So far I have argued that Robust Conservatism is false and Anemic Conservatism is consistent with Mooreanism. In this section I will argue that the same question that makes trouble for conservative opponents of Mooreanism—whether conservatism is a claim about propositional justification in the robust or anemic sense—also makes trouble for the most compelling argument in favor of conservatism. The argument I have in mind tries to show that both liberalism and Moorean reasoning are inconsistent with Bayesianism. This argument is well regarded by both liberals and conservatives. The latter see it as a reason to abandon liberalism,

the former as a reason to abandon orthodox Bayesianism: that is, to give up the view that epistemic probability, or rational credence, is governed synchronically by the norms of probabilistic coherence and diachronically by the claim that Pr(p) upon learning that E should equal one's prior Pr(p|E).[19] If this argument is sound, Mooreans face a puzzle. Fortunately, though, both sides have got it wrong by overstating what conclusions about propositional justification we can read off of the Bayesian formalism.

The argument rests on the following formal results showing that my confidence in an ordinary proposition on the basis of some evidence E cannot exceed the rational confidence that I had in some anti-underminer *prior* to learning that E. For example:

Pr(I am perceiving a hand right now|I am currently having an experience of a hand) ≤ Pr(I am not a BIV being deceived into thinking I am seeing a hand)

Pr(I know that it will be cold in Pittsburgh next January on the basis of induction|all previous Januaries in Pittsburgh have been cold) ≤ Pr(I am not in a world where the laws of meteorology change abruptly in 2010)

It is easy to see why these inequalities hold. First, my confidence that I am currently perceiving a hand cannot exceed my confidence that I am not a deceived BIV (since *I am a deceived BIV* entails *I am not perceiving a hand*). So for any body of evidence E,

Pr(I am perceiving a hand right now|E) ≤ Pr(I am not a BIV being deceived into thinking I am seeing a hand|E).

Second, having a perceptual experience as of a hand *increases*, rather than decreases, the epistemic probability of *I am a BIV being deceived into thinking I am seeing a hand*. Together, these claims imply that my confidence that I am perceiving a hand *after* having an experience as of a hand cannot exceed my *prior* confidence in *I am a BIV being deceived into thinking I am seeing a hand*. Analogous reasoning applies in the inductive case. On the basis of these sorts of results, it can seem like I can only have justification to believe that I am perceiving a hand right now on the basis that I am currently having an experience as of a hand if I have antecedent or

---

[19] The most fully-developed version of the Bayesian argument against liberalism appears in White; a quicker version appears in Schiffer. Pryor (in his unpublished manuscript "Uncertainty and Undermining") and Weatherson both acknowledge the force of the argument and accordingly urge revision of Bayesianism to accommodate liberalism. Silins also cites a similar argument which he uses to develop a view that mixes liberalism with rationalism.

independent justification to believe that I am not a brain-in-a-vat being deceived into having an experience of a hand.

But of course, we cannot read that result about *propositional justification* from some formal results about *epistemic probability*. In particular, we can only read conservatism off of these results given the following assumption:

Auxiliary Thesis about Independent Justification (or ATIJ): if someone believes that X on the basis of E, and the probability of Y prior to learning that E is an upper bound on the probability of X after learning that E, justification to believe that Y must be independent of justification to believe that X. [20]

Why believe ATIJ? ATIJ says that that if my credence in Y before acquiring evidence E is an upper bound on my credence in X after learning that E (where X is based on E), then if X is justified on the basis of E then I must have independent justification to believe that Y. The thought is that if X is justified on the basis of E, $Pr(X|E)$ must be high enough to make believing that X appropriate upon updating on E. Since $Pr(Y)$ is an upper bound on $Pr(X|E)$, the prior probability of Y is an upper bound on the probability of X upon learning E. So the prior probability of Y must be high enough to render belief in Y justified. Since Y had this degree of probability before X was justified, Y having this degree of probability cannot possibly depend on X being justified. So, justification to believe that Y must be independent of justification to believe that X.

ATIJ is a claim about propositional justification, so of course we should pause to ask: is it a claim about robust or anemic justification? This question yields a familiar dilemma. One the one hand, if ATIJ is a claim about *robust* propositional justification it is false. On the other hand, if ATIJ is a claim about *anemic* propositional justification it makes no trouble for Mooreans. To see why ATIJ is false if interpreted as a claim about robust propositional justification, consider a case involving reasoning-directed defeat:

The Failed Logician: Stella took, and failed, a logic class in college. She had particular difficulty mastering conditionals. One day Stella learns that E, which is in fact excellent evidence for p, and becomes extremely rationally confident that p. Yet she is rightly agnostic about all material conditionals, including E→p. She remembers that the rules governing conditionals are a bit tricky and counterintuitive, so she is rightly leery of committing to any.

---

[20] I owe this point to Pryor, who makes it in his unpublished manuscript "Uncertainty and Undermining".

The conditional E➔p is an upper bound on Stella's confidence in p, and $Pr(E➔p|E) \leq Pr(E➔p)$, so according to ATIJ Stella can have justification to believe that p on the basis of E only if she has p-independent justification to believe the conditional E➔p. Yet she lacks justification to believe that conditional in the robust sense—a reasoning-directed defeater prevents anyone with her evidence from engaging in sound reasoning concluding in E➔p—so, according to robust ATIJ, she cannot form a justified belief that p. But that's absurd: it cannot be that Stella's reasonable agnosticism about conditionals renders her incapable of acquiring any justified beliefs on the basis of any evidence!

Why does robust ATIJ fail? That is, why can't we read conclusions about robust propositional justification from Bayesian results? The problem is that Bayesianism abstracts from reasoning-directed evidence. So while Bayesianism reveals relations of object-directed evidence, it does not reveal much about robust propositional justification. That is why, for instance, Bayesianism says everyone should assign logical truths—even complex ones— probability 1. The evidence that makes it, in practice, irrational to be so confident in logical truths is reasoning-directed evidence: each of us knows that he makes logical errors from time to time.

The failure of Bayesianism to take reasoning-directed evidence into account has implications beyond the implausible assumption of logical omniscience. One of Bayesianism's two central claims is that we should update our credences via conditionalization: that is, $Pr(p)$ upon learning that E should equal one's prior $Pr(p|E)$. However, this rule of updating fails as a recipe for responding to reasoning-directed evidence. Consider Detective Smith again, and call the evidence in his case E. Before conducting his investigation, he might correctly reason that:

Pr(the butler did it|E)=Pr(the butler did it|E and Smith is an unreliable detective)

since whether Smith is an unreliable detective does not bear at all on whether the butler did it. Later, Smith learns that E, and becomes confident that indeed the butler did it. Later still, Smith learns that he is an unreliable detective and lowers, somewhat, his confidence that the butler did it. All of this seems appropriate. But note that Smith has violated the rule of updating his credences via conditionalization. That rule would require his confidence that the butler did it to remain unchanged when he learns that he is an unreliable detective. But Smith is surely right to

violate the rule: it would be unreasonable for him not to lower his confidence that the butler did it in light of his new evidence. So, reasoning-directed evidence generates counterexamples to the principle that one's credence in a proposition upon acquiring some total body of evidence should equal the prior epistemic probability of that proposition conditional upon having that total body of evidence.[21] The rule that you ought to update via conditionalization is plausible as a claim about how to respond to new object-directed evidence, it is a poor guide to how we ought to respond to new reasoning-directed evidence. Bayesianism is thus best thought of as a guide to object-directed support; it is just silent on reasoning-directed evidence. We cannot use Bayesianism as a guide to *robust* propositional justification.

So the robust version of ATIJ, like Robust Conservatism, yields obviously false conclusions in cases involving reasoning-directed defeat. What about anemic ATIJ? Anemic ATIJ may well be true – the Failed Logician is no counterexample to it, and indeed I am unable to think of any counterexamples. Moreover, the Bayesian results above show that someone can only acquire justified beliefs on the basis of perception if he already has adequate object-directed evidence to believe *I am not a deceived BIV*. So if we think of anemic propositional justification for p as having a lot of object-directed evidence for p, anemic ATIJ has got to be correct. As a result, all non-skeptics should learn to live with some anemic form of rationalism: it follows from the above Bayesian results and anemic ATIJ. But anemic ATIJ, combined with the Bayesian results above, will only yield an argument for Anemic Conservatism. As we've seen, that's no threat to Mooreanism. So the Bayesian considerations do not tell us what is wrong with Moore's proof. They do, however, suggest that the picture fleshed out in parts 4 through 6— wherein Moorean reasoning can be fine, but we might need to learn to live with anemic rationalism—is correct.

---

[21] David Christensen makes this argument in his unpublished manuscript "Higher-Order Evidence".

2.8   CONCLUSION: REASONING-DIRECTED DEFEAT AND AGNOSTICISM

The upshot of my argument is that the Achilles heel of the conservative response to Mooreanism lies in a surprising place: an ambiguity about whether having propositional justification entails the existence of a sound, undefeated bit of reasoning from one's evidence to the proposition in question. It's easy to see how this problem could remain invisible. It is tempting to assume that the following two claims are equivalent:

1. S has adequate evidence that bears on whether ~U

2. There is some sound, undefeated bit of reasoning from S's evidence to the conclusion that ~U.

But—crucially—this isn't quite right.   Whether there is a sound deliberative route from someone's evidence to ~U depends, in large part, upon her reasoning-directed evidence, whereas whether someone has adequate evidence that bears on whether ~U depends solely upon her object-directed evidence. These two ways of spelling out *propositional justification to believe that ~U* are not equivalent. Yet only when we conflate them is the conservative diagnosis of Mooreanism tenable: either disambiguation yields a version of conservatism that can't make trouble for Mooreans.

This approach to Mooreanism leaves two questions. First: I've identified a type of case—involving reasoning-directed defeat for anti-underminers—where all parties should agree that Moorean reasoning is a way to acquire justification. Should we think of these cases as a sideshow to the central debate over Mooreanism, or rather accept that the debate over Mooreanism turns out to be just a debate about reasoning-directed defeat? My answer is that we ought to take the latter position: the debate over reasoning-directed defeat is the main event. The first sort of reason to adopt this position is that it allows Mooreans to avoid biting any bullets. It explains, in a non-*ad hoc* manner, why Moorean reasoning in cases such as Moore 1-3 and Induction 1-3 is kosher while allegedly Moorean reasoning in odious cases such as Table 1-3 and Election 1-3 is not. We thus get a response to skeptics without any objectionable "easy justification". Moreover, it offers a neat reply to the Bayesian objection to Mooreanism that

turns on an independently recognized failure of Bayesianism to account for reasoning-directed defeat.

Thus the stance that Mooreanism *just is* a doctrine about the rational response to reasoning-directed defeat avoids the problems widely believed to tar all Moorean views. However, this stance gives rise to a second question: isn't this a really astonishing place for the conservative response to Mooreanism to falter? Alternately: how on earth could a debate about a pattern of anti-skeptical reasoning and a debate about the rational response to a type of defeater boil down to the same thing? Answering this question requires a close look at rational agnosticism. Conservatives and liberals, recall from part 2, disagree only about how to rationally respond to agnosticism about underminers: that much is old news. Most discussions of liberalism and conservatism proceed as if it's just obvious what sort of agnosticism is at issue. But it is quite difficult to articulate just what sort of agnosticism about underminers is relevant to the disagreement between liberals and conservatives - and impossible to do so without invoking reasoning-directed defeat.

Many possible senses of "agnosticism" are obviously not relevant to the disagreement between liberals and conservatives. One type of agnosticism is simply failing to form an attitude about a proposition because you've never thought about whether it is true. But everyone agrees that people do not need to have ocurrent beliefs that the relevant anti-underminers are false, so this sort of agnosticism cannot be the relevant kind. A second kind of agnosticism occurs when someone lacks the conceptual capacities to entertain a proposition – but this, too, is obviously not the kind of agnosticism at issue. A third kind of rational agnosticism is having a bunch of evidence both for and against a proposition, where the two sides have exactly equal weight. But that is properly called indifference rather than agnosticism, and in any case doesn't describe our position vis-à-vis skeptical hypotheses.

These familiar options do not exhaust the ways in which someone can be rationally agnostic about a proposition. Another way is to have reasoning-directed evidence which suggests that you are in no position to figure out whether the proposition is true. Consider Detective Smith again. He may have evidence that the butler did it, and he certainly has bothered to think about whether the butler did it, but he should be agnostic about whether the butler did it: not because he lacks sufficient object-directed evidence that the butler did it, but because he rationally ought to lack confidence in his ability to figure out the force of that

evidence. It's precisely this sort of rational agnosticism that generates the problem cases for conservatism: all of the counterexamples to Robust Conservatism in part 4 involve this kind of agnosticism.

This final type of agnosticism describes our position vis-à-vis skeptical scenarios, at least prior to engaging in Moorean reasoning: maybe we have adequate object-directed evidence to just write them off, but none of us ought to be very confident in his ability to figure out which deeply contingent propositions are true using reason alone.[22] After all, skeptics do not call our knowledge of the relevant anti-underminers into doubt by producing object-directed evidence against those anti-underminers – skeptics do not produce one iota of evidence that you are in fact a brain in a vat or that induction will suddenly become unreliable in the near future. They do not—or at least should not—argue that none of us has any evidence that skeptical hypotheses are false: maybe it is just impossible to run a computer powerful enough to deceive a brain into having experiences like mine, so my experiences are evidence that I am not a deceived brain in a vat. What skeptics do instead is call into question whether any of us is in a position to determine, on the basis of his evidence, whether those anti-underminers are true: maybe my evidence is inconsistent with my being a brain in a vat, but I am in no position to tell whether it is. Skeptics sow reasoning-directed defeat.

So it is no surprise that thinking carefully about reasoning-directed defeat reveals a problem with the conservative objection to Mooreanism: we cannot properly formulate what liberals and conservatives disagree about without discussing it. Moorean reasoning is nothing more than a tool for getting around agnosticism generated by reasoning-directed defeat. In our day-to-day lives, we do not need any such tools. When responding to skeptics, though, Moorean reasoning is just what we need.

---

[22] Is there another species of agnosticism: lacking *any* object-directed evidence, or more generally, adequate object-directed evidence, for or against a proposition? I think this is just a species of the sort of agnosticism I am discussing. How much object-directed evidence counts as *adequate* depends upon your reasoning-directed evidence: Detective Smith and Sherlock Holmes require different levels of object-directed evidence in order to have justification to believe that the butler did it. So this final type of agnosticism is not having enough object-directed evidence to clear the bar that your reasoning-directed evidence says you must clear in order to have justification to believe a proposition.

3.0   RATIONALIST RESPONSES TO SKEPTICISM: A NEW PUZZLE

Skeptical arguments focus on "bad" scenarios – situations where the world is not as we believe. There are lots and lots of ways for reality to fail to match my beliefs: why have a tiny handful of such "bad" scenarios achieved notoriety? For instance, I believe that I have hands. There are many "bad" scenarios where I don't, but some are a great deal more troubling than others. Compare a notorious "bad" hypothesis:

> **Brain in a Vat (BIV):** I am a recently envatted brain being fed deceptive sensations that mimic perception

with a rather less troubling one:

> **Stumps:** everything is as I believe it to be, except that there are no hands on the end of my arms

I don't have hands in either scenario, but only the first scenario is apt to generate a skeptical attack on my justification to believe that I have hands. Why? Roughly, because according to a skeptic armed with Stumps, perception works just fine – so I can trust my faculty of perception to tell me whether or not there are hands on the ends of my arms. BIV doesn't just say that I don't have hands, it also gives me reason to doubt that perception—the method I would normally use to check whether I have hands—can be trusted. BIV defeats the method I need to check whether I have hands. Similarly, compare the notorious:

> **Non-Uniform World (NUW):** the laws of nature will suddenly change tomorrow

with the tame:

> **Gloomy:** everything is as I normally believe it to be, except that the sun won't rise tomorrow.

If either hypothesis is true, I can't know that the sun will rise tomorrow. Gloomy will not help to generate a worrying skeptical attack, though. The reason why not is that it is easy to see how I can know that it doesn't obtain: easily accessible astronomical data and basic physics allow me to figure out that the sun will, indeed, rise tomorrow. Of course, the same astronomical data and physical laws entail that NUW is false, too. However NUW is more worrying because it entails that precisely the law-based inductive inference that I'd normally use to figure out that the sun will rise tomorrow can't be trusted. What separates the nasty hypothesis from the tame ones isn't that they entail the falsity of some of my ordinary beliefs (I have hands, the sun will rise tomorrow). Both nasty and tame hypotheses do that. Rather, what separates the nasty hypotheses is that they also entail that the sort of reasoning I would use to figure out whether *the hypotheses themselves* are true are untrustworthy across the board.

How should we respond to the notorious skeptical hypotheses? One way is to give in and say that we can't form justified beliefs that that BIV and NUW are false. That concessive response is unappealing, so let's set it aside for now. Non-concessive lines of response fall into two major camps: the division between the two camps centers on how, exactly, we ought to think of the relation between notorious and tame skeptical hypotheses.

The first camp says we ought to treat the notorious skeptical hypotheses *just like the tame ones*. How do I know that Stumps is false? I can see that I have hands – so there are hands, not stumps, on the end of my arms. How do I know that BIV is false? I can see that I have hands – so I'm an embodied person, not a brain-in-a-vat. Similarly, I have lots of inductive evidence that the laws of nature won't suddenly change tomorrow: they've always remained constant in the past, after all. Call responses in this camp **Moorean**, since they recall G. E. Moore's notorious "proof" of an external world. Moorean responses seem fishy: by relying on the very ways of acquiring justification that the skeptical scenarios call into question, these replies seem unresponsive or question-begging. They appear to ignore the plain fact that the notorious skeptical hypotheses are nastier than the tame ones.

The second non-concessive camp grants skeptics that our normal ways of forming beliefs about our surrounding or about the future won't cut any ice against the notorious hypotheses. But members of the second camp stop short of saying we have no way at all of forming justified beliefs that notorious skeptical hypotheses are false: they adopt the **rationalist** position that we can form justified beliefs apriori—that is, beliefs not based on any evidence—that skeptical

scenarios such as BIV and NUW are false. Indeed, many rationalists believe that the notorious hypotheses show what we must take for granted in order to be in the business of forming beliefs about the world – we can't reason our way to the conclusion that they are false because their falsity is a presupposition of reasoning at all.[23]

One very common sort of worry about rationalism is that it is *mysterious*: how could there be the sort of apriori justification that it requires? BIV and NUW are deeply contingent propositions, so the rationalist story requires deeply contingent apriori justification – many philosophers who accept apriori justification for beliefs about necessary claims will balk at apriori justification for beliefs about deeply contingent claims[24]. The latter sort of justification seems weird: how do we get it? Of course, anyone trying to land this attack will need to make it *a lot* more precise, but it is hard not to sympathize with the thought that we can't get justified beliefs in deeply contingent claims for free.

I'd like to focus, instead, on a second worry about rationalism: that it *just pushes the skeptical problem back*. A clever skeptic wielding BIV or NUW should, if faced with a rationalist reply, just trot out another skeptical hypothesis:

> **The Very Manipulative Neuroscientist (VMN):** I am a recently envatted brain and a neuroscientist is manipulating my faculties of reason in ways that make it seem to me that I have apriori justification to believe false, deeply contingent propositions for which I do not, in fact, have apriori justification.

VMN appears to stand toward apriori reasoning as BIV stands to perception and NUW stands to induction: it defeats the very method of forming beliefs that we'd need to use to show it is false. It's a nasty skeptical hypothesis. If a skeptic can use VMN to remove our apriori justification for deeply contingent propositions then he can use VMN to remove our apriori justification for not-BIV and not-NUW. An inability to respond to a skeptic wielding VMN renders rationalism impotent as an anti-skeptical strategy. What should a rationalist say about how anyone is able to acquire a justified belief in not-VMN? Saying that we can use apriori reasoning to determine

---

[23] For an eloquent defense of this position, see Wright (1985), (2002), (2004), and (2008) as well as White (2006).
[24] A proposition is superficially necessary if it is necessary because of the way that the reference of some term is fixed; it is superficially contingent if it is not superficially necessary. In contrast, a proposition is deeply contingent if there are some possible worlds in which a state of affairs that makes it true obtains and others where no such state of affairs obtains. The origin of this distinction is Evans (2000).

For a defense of deeply contingent apriori justification, see Hawthorne (2002).

that VMN is false looks non-responsive or question-begging in just the same way that using induction to argue that NUW is false or citing perceptual evidence that BIV is false seem non-responsive or question-begging. Rationalism thus seems to lose the big advantage it had over Mooreanism: its promise to refute skepticism without just begging the question against skeptics. Call this claim the *Same Boat Hypothesis*. The Same Boat Hypothesis creates a dilemma for fans of rationalism as an anti-skeptical strategy. If question-begging replies are no good, then a skeptic armed with VMN wins. If question-begging replies are okay, it's not clear why we need rationalism in the first place: we should become Mooreans and be done with it. Either way, rationalism fails as an anti-skeptical alternative to Mooreanism: either both strategies fail (so the skeptic wins) or rationalism can defeat skeptics but so can Mooreanism.

This second worry is far more damning than the first. The first worry doubts that we have the sort of apriori justification that rationalists insist on. The second worry says that even if we grant rationalists all that they ask—apriori justification to believe in the falsity of deeply contingent skeptical hypotheses—it isn't enough to defeat skeptics without making Moorean inferences. Rationalism, absent Moorean inference, doesn't work as an anti-skeptical strategy *even if it is true*. As far as I can tell, this second worry has received no attention in the literature on skepticism.

My aims in this paper are twofold: first, to determine whether rationalism and Mooreanism are in the same boat when it comes to circularity; second, to determine whether that boat is a bad one to be in. My comments so far have been imprecise: if we are going to get straight about whether there is a real problem for rationalism here, we will need to say a lot more about exactly what *question-begging* means in this context. The next part of this paper will present skeptical challenges centered on BIV, NUW, and VMN more carefully. Part II will pose the question of what, exactly, is viciously circular or question-begging about Moorean responses to skepticism. Parts III, IV, and V will lay out three possible ways to make this charge of circularity more precise. Each of these attempts will fail in two ways. First, each will fail to give a principled account of why the alleged constraint on reasoning that rules out Mooreanism does not also rule out the rationalist response to a skeptic wielding VMN – that is, each will entail the Same Boat Hypothesis. Second, each will fail to point out a plausible general constraint on reasoning that is violated by Moorean responses to skepticism – each will show that the boat shared by Mooreanism and rationalism is a fine one to be in. I conclude that the

dilemma posed above succeeds: rationalism and Mooreanism are in the same boat, so rationalism is not an anti-skeptical alternative to Mooreanism. Does this mean that rationalism is *false*? Certainly not; it just means that rationalists are in no position to condemn Moorean reasoning as unacceptably circular. This result defeats an influential argument in favor of rationalism—namely, that it's the only plausible alternative to skepticism—but does not give any affirmative reason to believe that rationalism is false. Indeed, I am not at all sure that anyone—Mooreans included—should reject rationalism entirely. In the final part of the paper I'll argue that we shouldn't think of Mooreanism and rationalism as rivals, but rather as compatible—and possibly complementary—theories about different things.

3.1 RESPONSES TO SKEPTICISM

How, exactly, is a skeptic supposed to use BIV to threaten our ordinary perceptual knowledge? A common way to present the skeptic's argument is:

> **Premise:** (Simple-1) I do not have justification to believe that not-BIV
> **Premise:** (Simple-2) If I do not have justification to believe that not-BIV then I do not have justification to believe that I have hands
> **Therefore:** (Simple-3) I do not have justification to believe that I have hands

It is not clear that anyone should lose sleep over this argument. The first premise is itself a skeptical conclusion, and hardly something that most people accept pre-philosophically. So it looks like we are free to reject it without further ado. Yet without Simple-1, skeptics cannot make any mischief with this style of argument.

Skeptics can do better. A stronger argument in the same neighborhood—one that doesn't assert that I lack justification to believe that not-BIV by fiat—is available:

Premise: (Nasty-1) Either I do not have justification to believe that not-BIV *or* I do have justification to believe it, but my justification depends (at least in part) on beliefs justified by perception

Premise: (Nasty-2) If I have justification to believe any proposition on the basis of perception, then I must have independent (of perception) justification to believe that not-BIV

So: (Nasty-3) I do not have justification to believe that not-BIV

Premise: (Nasty-4) If I do not have justification to believe that not-BIV then I do not have justification to believe that I have hands

Therefore: (Nasty-5) I do not have justification to believe that I have hands

The third step of this argument follows from the first two because Nasty-1 says that any justification I have to believe that not-BIV depends upon perception, while Nasty-2 says that justification to believe *anything* on the basis of perception requires perception-independent justification to believe that not-BIV. More simply, Nasty-1 says that justification to believe that not-BIV could *only* come from perception, while Nasty-2 says that justification to believe that not-BIV *cannot* come from perception. Nasty-3 follows because there is nowhere left for justification for not-BIV to come from.[25]

It's worth saying a bit about the notion of *dependence* that figures in this argument. What I have in mind here is not a kind of psychological dependence—this argument does not turn on a claim about the psychological process wherein someone forms a belief that not-BIV—but rather on rational dependence. A belief that p rationally depends on perception just in case were I to learn that my faculties of perception were totally unreliable, I would lose my justification to believe that p.

The situation regarding inductive skepticism is analogous. Consider a familiar sort of argument:

Premise: (Simple Induction-1) I do not have justification to believe not-NUW

Premise: (Simple Induction-2) If I do not have justification to believe that not-NUW then I do not have justification to believe that the sun will rise tomorrow

Therefore: (Simple Induction-3) I do not have justification to believe that the sun will rise tomorrow

---

[25] My discussion of the superiority of Nasty 1-5 to Simple 1-3 follows Pryor (2000).

Non-skeptics can comfortably deny Simple Induction-1, so this argument isn't very powerful. The more worrying argument is:

> **Premise:** (Nasty Induction-1) Either I do not have justification to believe that not-NUW *or* I do have justification to believe it, but my justification depends (at least in part) on beliefs justified by induction
> **Premise:** (Nasty Induction-2) If I have justification to believe any proposition about the future on the basis of induction, then I must have independent (of induction) justification to believe that not-NUW
> **So:** (Nasty Induction-3) I do not have justification to believe that not-NUW
> **Premise:** (Nasty Induction-4) If I do not have justification to believe that not-NUW then I do not have justification to believe that the sun will rise tomorrow.
> **Therefore:** (Nasty-5) I do not have justification to believe that the sun will rise tomorrow

Where, again, the relevant dependence is *rational* dependence. There is no premise here that non-skeptics can flatly reject, so once again the nasty argument is far more effective than its simple counterpart.

How might we resist Nasty 1-5 and Nasty-Induction 1-5? Each argument has three premises, and thus three possible loci of resistance. One way out is to deny the fourth step of each argument, and say e.g. that I do not have justification to believe that not-BIV but do have justification to believe that I have hands. The trouble here is that brains in vats don't have hands, so *I have hands* entails not-BIV. Thus this strategy involves denying **closure**—the thesis that if someone has justification to believe that p, and p entails q, and he forms a belief that q based on p, then his belief that q must be justified—since if closure were true, anyone with justification to believe that he has hands would have justification to believe not-BIV. Denying closure is a radical move indeed, so from now on I'll proceed on the assumption that closure is true.[26]

A second locus of resistance is the first premise of each argument, which says that *if* I have justification to believe not-BIV, that justification depends upon perceptual beliefs, and *if* I have justification to believe not-NUW, that justification depends upon inductive beliefs. These premises do not say that the justification in question must come from perception or induction *alone*: rather, they claim that knowledge of my surroundings must depend, at one stage or another, on beliefs justified by perception and that knowledge of the future must depend, at one

---

[26] An adequate defense of closure would take us far afield. For an argument for closure, see Hawthorne (2005).

stage or another, on beliefs justified by induction. Even so, rationalists deny both of these premises since they insist that we have apriori justification for not-BIV and not-NUW – and this apriori justification in no way depends upon perceptual or inductive beliefs. So rationalists can resist the first premise of each argument, but everyone else looks stuck with it.

The final locus of resistance is the second premise of each nasty argument: the claim that using perception to acquire justification requires independent (of perception) justification to believe not-BIV, and using induction to acquire justification requires independent (of induction) justification to believe not-NUW. This premise is an anti-circularity requirement: it says you cannot acquire justification to believe anything on the basis of perception unless you have justification to believe not-BIV—that much follows from closure—and that, moreover, you cannot acquire justification to believe not-BIV for the first time by using perception: that is, you cannot reason *here is a hand (I perceive as much) therefore not-BIV*; Nasty Induction-2 rules out the analogous inductive case for not-NUW. The second premise of each argument, then, rules out the sort of circularity involved in Moorean reasoning.

We can now state our worry about rationalism more carefully. Suppose a rationalist rejects Mooreanism—and thus embraces Nasty-2 and Nasty Induction-2—and also embraces closure, and thus embraces Nasty-4 and Nasty Induction-4. That rationalist looks ill-positioned to respond to:

> **Premise:** (Manipulative-1) Either I do not have justification to believe that not-VMN *or* I do have justification to believe it, but my justification is depends (at least in part) on beliefs justified by perception *or* I do have justification to believe it, but my justification comes from apriori reasoning
> **Premise:** (Manipulative-2a) If I have justification to believe any proposition on the basis of perception, then I must have independent (of perception) justification to believe that not-VMN
> **Premise:** (Manipulative-2b): If I have justification to believe anything on the basis of apriori reasoning, then I must have independent (of apriori reasoning) justification to believe that not-VMN
> **So:** (Manipulative-3) I do not have justification to believe that not-VMN
> **Premise:** (Manipulative-4) If I do not have justification to believe that not-VMN then I do not have justification to believe that I have hands
> _____
> **Therefore:** (Manipulative-5) I do not have justification to believe that I have hands

Rationalism gives no grounds to reject the first premise, anti-circularity considerations seem to require accepting premises 2a and 2b, and closure requires accepting premise 4. So, rationalists

face a problem here. The only way out is to treat Manipulative-2b as different from all of the other second premises: that is, to say that while the Moorean response to BIV and the inductive response to NUW are viciously circular, using apriori reasoning to defeat VMN is acceptably circular. The Same Boat Hypothesis alleges the opposite, namely that any grounds for rejecting Manipulative-2b will also be grounds for rejecting Nasty-2 and Nasty Induction-2.

Some rationalists may question the first premises by insisting that justification to believe not-VMN is had by **default** and thus does not depend on anything, including apriori reasoning. If default justification is justification that does not require one to have gone through any psychological process of reasoning, then the claim that not-VMN has default justification doesn't matter much for our purposes. Nothing in Manipulative 1-5 turns on the process whereby anyone comes to believe not-VMN: the argument turns instead on claims about rational dependence. However, if default justification means justification that cannot be undercut—that is, justification such that there is no claim U such that, were I to learn U, my justification would diminish—then justification to believe not-VMN cannot be default. Imagine I came to know that I had been brainwashed: some nefarious experimental philosophers, as part of their research, deliberately and extensively tampered with my faculties of apriori reasoning (including my reasoning about which propositions are apriori). Learning *that* really ought to diminish my confidence in putative apriori claims, including (if I believe it to be apriori) not-VMN. So not-VMN cannot be default in the second, stronger sense: justification to believe it requires a shakable trust in my faculties of apriori reasoning. So insisting on default justification won't derail the first premise in any significant way.

Getting straight on whether rationalists can reject Manipulative-2b without rejecting all of the other second premises, though, will require a closer examination of just what the anti-circularity constraint involves. So far we've just gestured roughly at some constraint that rules out Moorean inferences such as justifying not-BIV by using perceptual beliefs. We're going to need to pin this constraint down precisely before we can figure out whether it makes trouble for rationalists faced with VMN as well.

3.2 WHY NOT BE A MOOREAN?

What is wrong with Moorean reasoning? So far we have gestured in the direction of a problem by saying that Moorean reasoning is circular or question-begging, but this is far too imprecise. Consider:

> **Enabling condition[27]:** (Moore-1) [an experience as of a hand]
> **So:** (Moore-2) I have a hand
> ---
> **Therefore:** (Moore-3) not-BIV

> **Premise:** (Moorean Induction-1) In the past, observed regularities have tended to continue into the future
> **So:** (Moorean Induction-2) From now on, observed regularities will tend to continue into the future
> ---
> **Therefore:** (Moorean Induction-3) not-NUW

Moorean reasoning does not beg the question in the typical sense of being **tautological**: the conclusions of Moorean reasoning are not among the premises. So circularity in *that* sense cannot be what is wrong with these bits of reasoning. None of this is to deny that these pieces of reasoning seem fishy, but what, exactly, is wrong with them?

There are a few other bad diagnoses. A first bad diagnosis is to deny the first step of each argument, i.e. to say that in the past observed regularities have *not* tended to continue into the future, or to say that nobody ever does have a perceptual experience as of a hand. That reply lacks any plausibility. A second bad reply is to deny that any of us ever has justification to believe the second step of each argument on the basis of the first. But that's an awfully drastic measure: we'd like a diagnosis of these arguments that does not commit us to wholesale skepticism about perceptual beliefs. Third, we could deny that anyone can infer the conclusion from the second step. However, in each case the second step deductively entails the conclusion, so this move requires denying closure; that, too, is an unappealing option. Can't we do better?

---

[27] I say "enabling condition" rather than "premise" since having some perceptual experience cannot itself be a premise in an argument (though, of course, the proposition *I am having a perceptual experience as of a hand* could be). By *enabling condition* I simply mean a state that licenses the formation of another belief, here via the use of perception.

Yes – in the next three sections I will look at three accounts that have some go in them. All three of these accounts will allege that Moore 1-3 and Moorean Induction 1-3 are counterexamples to **transmission**: the thesis that if you have justification to believe the premises of a bit of reasoning (or: are in fact in the enabling conditions for that bit of reasoning) and the premises or enabling conditions entail the conclusion, you can acquire additional justification to believe the conclusion by going through the bit of reasoning. That is, the diagnosis will say that someone cannot acquire additional justification—or, justification for the first time—to believe not-BIV by going through Moore 1-3, nor can someone acquire justification to believe not-NUW by going through Moorean Induction 1-3. These bits of reasoning exhibit *transmission failure*.

If the Same Boat Hypothesis is false, two things need to be true. First, there must be some plausible constraint on reasoning that entails Nasty-2 and Nasty Induction-2; in order to do this, it must entail that Moore 1-3 and Moorean Induction 1-3 cannot possibly give one justification to believe their conclusions. That is, we need a story about why Moore 1-3 and Moorean Induction 1-3 exhibit transmission failure. Second, this constraint on reasoning must not entail Manipulative-2b, on pain of defeating both Mooreanism *and* rationalism as anti-skeptical strategies. Obviously this second task is easy if we allow rationalists to put some *ad hoc* provisos in the relevant diagnosis: most egregiously, we can always attach a rider "*...except when the skeptical scenario is VMN*" to our constraint and get rationalism out of trouble. The challenge is rather to motivate a constraint wherein there is a plausible, non-*ad hoc* story about why Moorean reasoning is circular and rationalist responses to VMN are not that entails Nasty-2 and Nasty Induction-2 but not Manipulative-2a.

3.3   FIRST DIAGNOSIS: INSENSITIVITY

The first diagnosis begins with a simple observation: were BIV true, I would still have experiences as of hands. Similarly, were NUW true, induction would still have a good track record. Thus I could go through Moore 1-3 or Moorean Induction 1-3 even in "bad" worlds

where skeptical hypotheses hold. This seems to make Moore 1-3 and Moorean Induction 1-3 suspect at best. After all, we wouldn't trust a DNA test if we knew that it would say some blood was Smith's even if it wasn't: the *possibility* of a false positive might be tolerable, but *certainty* of a false positive would not! Yet that's just what Moore 1-3 and Moorean Induction 1-3 give us: it is certain that, were I in a "bad" world, I would be able to go through the reasoning Moore 1-3 or Moorean Induction 1-3 and thereby arrive at a false conclusion.

A little more carefully: call S's belief that p **sensitive** iff were it the case that ~p, S would not believe that p (and a belief is **insensitive** if it is not sensitive). A tempting position is:

> **The sensitivity constraint on justification:** If S knows that her belief that p is not sensitive, then her belief that p is not justified.

Note that this position is rather less ambitious than the claim that S knows that p just in case her belief that p is sensitive: it says only that *known* insensitivity is inconsistent with justification. If the sensitivity constraint on justification is true, any reasoning that we know to reach an insensitive conclusion without starting from any insensitive premises must fail to transmit justification to its conclusion. Yet Moore 1-3 and Moorean Induction 1-3 have just this feature. Moore 1-3 has *no* premises—and thus no insensitive premises—whereas Moorean Induction 1-3 has only the premise *in the past, observed regularities have tended to continue into the future*, and that belief *is* sensitive: were it false, I would not believe that it was true. Yet we know that the conclusions of both arguments are insensitive. Thus, we can condemn both Moorean arguments as failing to transmit justification. Call this the **insensitivity diagnosis.**

The insensitivity diagnosis turns on three features of skeptical arguments:

1. Our ordinary, humdrum beliefs tend to be sensitive: if I didn't have hands, I wouldn't believe that I did and if the past had been wildly irregular, I wouldn't believe that it was mostly regular.

2. Our beliefs in the negations of notorious skeptical hypotheses such as BIV and NUW are known to be insensitive. In contrast, our beliefs in the negations of tame skeptical hypotheses, not-Stumps and not-Gloomy, are sensitive. If I had stumps on the end of my arms, I'd see that I didn't have hands; if astronomical laws gave reasons to think the sun wouldn't rise tomorrow, it'd be all over the news.

3. There are entailment relations between ordinary propositions and skeptical hypotheses. For instance, if I have hands then not-BIV. Sensitivity is not *closed* under entailment: p can entail q, where p is sensitive, and yet q is insensitive.

These three features create a puzzle: most of us want to accept closure, yet most of us (initially, anyway) find something like the sensitivity constraint on justification plausible. It's hard to know how to resolve this tension.

This puzzle accounts for a lot of our schizophrenia about skeptical arguments. Mooreans resolve this tension one way: our (sensitive) ordinary beliefs are justified, and our ordinary beliefs entail the falsity of skeptical hypotheses, so by closure we can form justified beliefs in the falsity of skeptical hypotheses. Skeptics resolve it the other way: our (obviously insensitive) beliefs in the falsity of skeptical hypotheses are unjustified, yet our ordinary beliefs entail the falsity of those skeptical hypotheses, so by closure our ordinary beliefs must not be justified. A third way to resolve this tension is, of course, to keep the sensitivity constraint and just give up closure: this allows maintaining both that our ordinary beliefs are justified and that our beliefs in the negations of skeptical hypotheses are not.[28]

As an account of our intuitions about skeptical arguments, the insensitivity diagnosis has a lot going for it. However, fans of the Same Boat Hypothesis cannot embrace the insensitivity diagnosis of what's wrong with Mooreanism for two reasons. First, it requires the sensitivity constraint on justification, yet as we have seen that account is hard to square with closure. Giving up closure is a very high price to pay. Second, though, the insensitivity diagnosis won't give us what we are after: a position that rejects the Same Boat Hypothesis. According to the insensitivity diagnosis, the rationalist response to VMN is circular *in just the same way* as Moorean reasoning. Belief in not-VMN is insensitive, so the insensitivity diagnosis condemns any reasoning that alleges to lead to a justified belief in VMN. Put another way, since apriori reasoning to not-VMN has no insensitive premises but an insensitive conclusion, the insensitivity diagnosis must condemn it as bad reasoning. Thus, fans of the insensitivity diagnosis are committed to Manipulative-2b. The insensitivity diagnosis does not pick out a kind of circularity

---

[28] Nozick (1981) defends the sensitivity account of knowledge and adopts this third strategy (adapted to focus on knowledge, rather than justification): that is, he denies closure and thus claims that he knows that he has hands but does not know that not-BIV.

involved in Moorean arguments that does not also appear in the rationalist reply to VMN. Our first attempt has failed.

3.4   SECOND DIAGNOSIS: A DIALECTICAL CONSTRAINT

Most people are inclined to say that Moorean reasoning "begs the question" against skeptics, though as we have seen Moorean arguments, unlike other "question-begging" arguments, aren't tautological. However, etymologically "begging the question" has nothing to do with tautologies – rather, it refers to arguing in a way that just asks (i.e. begs) one's opponent to grant one's conclusion, rather than *earning* the conclusion. It is thus a dialectical no-no: a question-begging argument should not move someone who disagrees with you, and is thus unwilling to just grant your conclusion. A second diagnosis of Moorean reasoning is that it commits just this sort of dialectical foul. A little more carefully, we might embrace:

> **No Question-Begging:** A belief that p is justified only if it rests upon reasoning that would have some dialectical weight against someone who believes that ~p.

Where an argument has *dialectical weight* against someone if and only if it provides him with *some* reason to believe its conclusion. It is easy enough to see why No Question-Begging makes trouble for Moorean reasoning. Moore 1-3 has no dialectical weight against someone who believes BIV: he'll deny, with perfect internal consistency, that Moore-1 provides any support for Moore-2. Similarly, Moorean Induction 1-3 has no dialectical weight against someone who believes NUW. Call this account of the problem with Moorean reasoning the **dialectical inefficacy diagnosis**.

The dialectical inefficacy diagnosis is not new. Jim Pryor cites it as the reason why Moorean reasoning *seems* to be bad reasoning.[29] Pryor grants that Moorean reasoning is

---

[29] See Pryor (2004).

dialectically ineffective, and thinks that we all feel some intuitive pull towards No Question Begging. However, he thinks No Question Begging is false: it relies on a conflation between the task of defending a position against a skeptic and justifying your beliefs to yourself. Thus, No Question Begging can explain the sociological datum that a lot of people are uneasy with Moorean reasoning but cannot give a successful account wherein Moorean reasoning does not lend justification to its conclusions. Of course, just pointing out that defending a belief against a skeptic and justifying it to oneself are different tasks is not enough to defeat No Question Begging, since it is plausible that the criteria for success in those two tasks cannot come very far apart. A natural way to think of justifying a belief is to defend it against a possible skeptic in an imaginary dialogue. If that's right, No Question Begging does not rely on sloppy conflation, but rather an appealing view wherein justification to oneself and dialectical efficacy against others cannot come apart. We cannot write the dialectical inefficacy diagnosis off without a fight.

So the good news for rationalist anti-Mooreans is that the dialectical inefficacy diagnosis does indeed condemn Moorean arguments and has quite a bit of initial plausibility. Moreover, while No Question Begging is similar to conservatism, it only comes into play when responding to skeptics: it thus does not lead to the quotidian counterexamples that make trouble for conservatism. The bad news, though, is that it also spells trouble for rationalism. The reason why is clear enough: nothing a rationalist can say will have any dialectical force against someone convinced of VMN.[30] In particular, any claim that not-VMN is apriori will lack dialectical force. So the dialectical inefficacy account entails Manipulative-2b, and thus that Manipulative 1-5 is a sound argument. The dialectical inefficacy diagnosis entails the Same Boat Hypothesis.

This leaves us with a puzzle: the dialectical inefficacy diagnosis is *extremely* plausible, but it rules out all anti-skeptical strategies. We'd better have a good reason for rejecting it. One such reason is obvious: given a choice between accepting No Question Begging and skepticism, and rejecting No Question Begging and skepticism, the latter certainly seems the more appealing package. While that reason may be *convincing*, it isn't very *enlightening*: the case for No Question Begging seemed plausible enough, and it is natural to wonder where it went wrong. We need a plausible story about why the dialectical inefficacy diagnosis is false.

---

[30] This character—the fellow convinced of VMN—is not a familiar one in the philosophical bestiary. However, he needn't be very different from the more familiar (from the epistemology literature, anyway!) character who believes BIV. The believer in VMN simply adds that it may seem as if it is apriori that not-BIV, but that this rational appearance cannot be trusted since, like perceptual appearances, it is the result of a deception.

Fortunately, we can get one. No Question Begging models justification on argument against an opposing position, but skeptics of the sort who push BIV, NUW, or VMN aren't really defending a position – at least, not if No Question Begging is correct. Skeptics do not produce evidence or argument for BIV, NUW, or VMN. No skeptic has any credible evidence that any of these bad scenarios hold. That's familiar enough: skeptical arguments do not proceed by providing reasons to believe skeptical hypotheses; they proceed by saying, first, that we lack conclusive grounds for believing the negation of skeptical hypotheses and second, that we therefore ought to be agnostic about them. The problem is that if No Question Begging is true, and skeptics have no dialectically effective arguments in support of BIV, NUW, or VMN, then skeptics necessarily lack justification to believe those claims. A skeptic who embraces BIV, NUW, or VMN is thus thereby unreasonable if No Question Begging is true: his belief lacks justification. So if we imagine ourselves in an argument against such a skeptic, we ought to be able to shake him from his position just by mentioning No Question Begging! If No Question Begging is right, reasonable skeptics embracing BIV, NUW, or VMN cannot exist.[31]

This last point is important because if we are arguing with a skeptic who believes BIV, NUW, or VMN then we can, by invoking No Question Begging, get him to be agnostic about his favored skeptical hypothesis. The question of whether, say, Moore 1-3 is good reasoning then becomes, according to No Question Begging: should Moore 1-3 make someone who is *agnostic* about BIV have justification to believe not-BIV? But that's just the same as our original question, namely: is Moore 1-3 good reasoning? No Question-Begging thus gets us no closer to figuring out whether Moore 1-3 is good reasoning. We can only make sense of dialectical weight in the context of argument against someone with an internally consistent position – yet if No Question-Begging is true then nobody with an internally consistent position can believe BIV, NUW, or VMN. Once we've dislodged our dialectical opponent from his belief in BIV, NUW, or VMN can we then go on to give a Moorean argument? That's the question we've been worried about throughout, but No Question-Begging doesn't look well-positioned to answer it. So the dialectical inefficacy diagnosis fails to tell us what is wrong with Moore 1-3 or Moorean Induction 1-3. All it tells us is something we already knew: *if* there is a story about why those

---

[31] Of course, its possible that the skeptic with whom we are arguing does not accept No Question Begging. Maybe, then, it is not unreasonable for him to hold not-BIV. However, the point remains that we can condemn Moorean reasoning using No Question Begging only if we apply that constraint to ourselves but not to our skeptical opponent – and I just can't see any motivation for this double standard.

bits of reasoning are no good, then those bits of reasoning are no good. The charge of dialectical inefficacy adds nothing.

3.5   THIRD DIAGNOSIS: CONSERVATISM

Let's set skepticism aside, for a moment, and consider some bad arguments about everyday matters:

> **Premise:** (Table-1) This table looks red
> **So:** (Table-2) This table is red
>
> **Therefore: (**Table-3) This is not a white table under red lights
>
> **Premise:** (Election-1) Someone just checked a box on a piece of paper and slid it into a slot
> **So:** (Election-2) Someone just voted
>
> **Therefore: (**Election-3) This is not just a rehearsal of an election

These look like terrible bits of reasoning. It should not be so easy to acquire justified beliefs in Table-3 or Election-3; if these arguments are not epistemically useless they generate odious "easy justification" for their conclusions.[32] Yet, as before, it can be tricky to say what is wrong with each bit of reasoning: neither is tautological, and each step looks individually unobjectionable. A plausible diagnosis is that the premise of each argument gives reason to believe the second step *only on the assumption that the conclusion is true*. That the table looks red gives you reason to believe that it is red only on the assumption that the lights aren't red; that someone put some paper in a ballot box gives you reason to believe that someone voted only on the assumption that there is an election going on. A little more carefully, the negation of the conclusion of each argument is an **underminer** for the support lent by the first step to the second: a claim that, were you to learn it, would remove whatever rational force the first premise

---

[32] Cohen (2002) and Wright (2008) offer versions of this worry about Moorean reasoning; Cohen discusses easy knowledge, rather than easy justification, but the basic shape of the worry is the same.

lends to the second step of each argument. Do you have justification to assume the conclusion of each argument? Maybe yes, maybe no: if you're in Ikea, it's a safe bet that the lights are white, but if you're in a fun house then all bets are off. So if you're in Ikea, you can reasonably go through the reasoning of Table 1-3, but if you're in a fun house you may not. Either way, you can rationally go through the steps of Table 1-3 only if you *already* have justification to believe Table-3; similarly, you can only rationally go through the steps of Election 1-3 if you *already* have justification to believe Election-3. So while these arguments aren't tautological, they are epistemically useless in just the same way that tautological arguments are. Nobody can use them to acquire justification to believe anything new.

We can generalize this account of the problem with Table 1-3 and Election 1-3. Our diagnosis rested, implicitly, on a principle called **conservatism**: the claim that someone can acquire a justified belief that p on the basis of some grounds G only if he has antecedent justification to believe that all hypotheses that undermine the support lent by that consideration to p are false. Conservatism says that someone has justification to believe Table-2 on the basis of Table-1 only if he already has justification to believe Table-3 (ditto for Election 1-3). If conservatism is true, Moore 1-3 and Moorean Induction 1-3 are epistemically useless for just the same reason that Table 1-3 and Election 1-3 are. A perceptual experience as of a hand gives me reason to believe I have a hand only if I may reasonably assume that not-BIV; induction having a good track record gives me reason to believe that induction will continue to be reliable only if I may reasonably assume not-NUW. This diagnosis is appealing not only because conservatism looks compelling on its own, but because it seems to offer the right diagnosis of humdrum bad reasoning like Table 1-3 and Election 1-3. Call this the **conservative diagnosis**.

Does conservatism make the world safe for anti-Moorean rationalists: that is, does it entail Nasty-2, Nasty Induction-2, and Manipulative-2a, but *not* entail Manipulative-2b?

The answer isn't obvious. Conservatism says I can form justified beliefs using perception only if I have antecedent justification to believe not-BIV and that I can form justified beliefs using induction only if I have antecedent justification to believe not-NUW. Does it also say I can form justified beliefs using apriori reasoning only if I have antecedent justification to believe not-VMN? If so, is there any reason to think that it doesn't entail Manipulative-2b? I can think of two possible reasons to think that conservatism doesn't entail the Same Boat Hypothesis:

1. **Restrict our notion of *undermining*.** Perhaps evidential support can be undermined, but apriori rational support cannot. If we understand the term "underminer" narrowly to mean a defeater than removes the support lent by some *evidence* to p, conservatism is silent on whether someone needs antecedent justification to believe not-VMN to form justified apriori beliefs.

2. **Restrict our notion of *justification to believe*.** We can distinguish stronger and weaker notions of *justification to believe*. On a **robust** understanding someone has justification to believe that p only if there is a sound, undefeated deliberative route from his evidence concluding in a doxastically justified belief that p. On that robust understanding, conservatism entails Manipulative-2b: it says someone can form justified beliefs on the basis of apriori reasoning only if he *already* has an undefeated sound deliberative route concluding in a justified belief that not-VMN. On a more **anemic** understanding, justification to believe that p only requires adequate evidence that p. Since we're granting rationalists that we don't need any evidence to believe not-VMN, we've all got adequate evidence to believe that not-VMN and this in no way depends on any evidence or reasoning. If conservatism is a claim about justification in the anemic sense—that is, if conservatism says you need adequate evidential support for ~U (where U is an underminer), but do not need to be able to reason your way to the conclusion that ~U, in order to form a justified belief that p on the basis of some grounds G—then conservatism does *not* entail Manipulative-2b.

So there are potentially two ways out for rationalist conservatives. In the remainder of this section I'll argue that the first is no good but the second is tenable – and, indeed, advisable for reasons having nothing to do with the Same Boat Hypothesis.

About strategy (1): restricting our notion of *undermining* allows far too much to count as good reasoning. It's often possible to use broadly Moorean patterns of reasoning in apriori contexts – if we restrict conservatism to the narrow conception of undermining then it won't tell us what's wrong with apriori forms of Mooreanism. Consider these conditional proofs:

Assumption for conditional proof: (Apriori Election-1) Someone just checked a box on a piece of paper and slid it into a slot
> **So:** (Apriori Election-2) Someone just voted
> **So:** (Apriori Election-3) An election is happening, rather than a rehearsal

Therefore: (Apriori Election-4) If someone just checked a box on a piece of paper and slid it into a slot, then an election is happening, rather than a rehearsal

Assumption for conditional proof: (Apriori Moorean Induction-1) In the past, observed regularities have tended to continue into the future
> **So:** (Apriori Moorean Induction-2) From now on, observed regularities will tend to continue into the future
> **So:** (Apriori Moorean Induction-3): not-NUW

Therefore: (Apriori Moorean Induction-3) If in the past, observed regularities have tended to continue into the future, then not-NUW

These bits of reasoning don't look any better than their aposteriori counterparts, so it's hard to see the appeal of an understanding of conservatism that condemns the aposteriori versions of these arguments while remaining silent on the apriori ones. Worse still, we get the result that someone *cannot* form a justified belief in not-NUW via Moorean Induction 1-3, but *can* form a justified belief in not-NUW by reasoning through Apriori Moorean Induction 1-4, then noting that observed regularities have held in the past, and then concluding that not-NUW.[33] Will it help to say that, though these bits of reasoning are apriori, the proposition in the assumption is still, in some sense, evidence for the second step? Not really; we can just tweak the argument further:

Assumption for conditional proof: (Rational Moorean Induction-1) In the past, observed regularities have tended to continue into the future
> **So:** (Rational Moorean Induction-2) It is rational to believe that, from now │on, observed regularities will tend to continue into the future
> **So:** (Apriori Moorean Induction-3): It is rational to believe that not-NUW

Therefore: (Rational Moorean Induction-3) If in the past, observed regularities have tended to continue into the future, then it is rational to believe that not-NUW

Here the assumption supports the second step, but it really can't be described as *evidence* for it. While the conclusion of this argument is different from the ones above—it will only allow us to derive a claim about rational belief—it is enough to keep skepticism at bay. So restricting our notion of undermining won't really block Moorean responses to skepticism; it only requires tweaking them slightly so that they involve conditional proofs. The first way out is no good.

---

[33] Of course, foes of Mooreanism can get around this result by arguing that the proposed reasoning exhibits transmission failure. But the version of conservatism we are now considering—one with a restricted notion of "underminer"—cannot explain why this reasoning would exhibit transmission failure. So we'd need some other principle for restricting the transmission of justification – and, of course, we'd need to be sure that principle didn't defeat any justification to believe not-VMN.

What about the second – weakening our notion of antecedent justification to believe so that acquiring a justified belief that p on the basis of G doesn't require being able to rationally conclude that underminers for the support lent by G to p are false ?   It turns out that, quite apart from considerations about VMN, conservatives ought to adopt the anemic notion of justification to believe.   If we plug the robust version of justification back into conservatism, it delivers implausible results in everyday cases. Consider:

The Seasoned Interviewer: Sasha has interviewed many criminal suspects for the police, and she is very good at her job. After much practice and study, she is excellent at telling truth from lies and she knows it. One day a colleague asks Sasha if a suspect who frequently touches his nose during an interview is more likely to be lying. Sasha replies, truthfully, that she isn't sure; she finds it very hard to articulate such rules in the abstract since her previous attempts to do so have been unsuccessful. Shortly thereafter, Sasha interviews a suspect who frequently touches his nose. Sasha acquires a justified belief that the suspect is lying, and her belief is based, in large part, on his frequent nose-touching.

This sort of example shows that it is possible to form a justified belief in p on the basis of G while having evidence inconsistent with being able to form a justified belief in the falsity of an underminer for the support lent by G to p. The reason Sasha cannot form a justified belief that the underminer is false isn't that she *lacks evidence* for the anti-underminer; rather, she has evidence that she cannot reason her way towards a justified belief in it. Yet anemic conservatism isn't hopelessly weak: it will show what is wrong with Table 1-3 and Election 1-3, on the assumption that there's no impediment to directly figuring out whether there's a rehearsal going on or what color the lights are.

All of this looks like good news for conservative rationalists: anemic conservatism won't entail manipulative-2b, so it won't make it impossible to respond to a skeptic armed with VMN. And anemic conservatism is plausible: it shows what's wrong with some obviously bad arguments. One question remains: will it block Moore 1-3 and Moorean Induction 1-3?

3.6   ANEMIC CONSERVATISM AND THE SAME BOAT HYPOTHESIS

Anemic conservatism doesn't threaten to leave rationalists unable to respond to a skeptic armed with VMN. However, adopting anemic conservatism only yields a way out of the Same Boat Hypothesis if anemic conservatism is inconsistent with Mooreanism. The problem is: it's not. Anemic Conservatism is consistent with Moorean reasoning sometimes being good reasoning. The problem cases arise when someone is in a position where (for some ordinary claim p based on grounds G where U undermines the support lent by G to p):

(a) He has p-independent justification to believe that ~U in the anemic sense

(b) There is no p-independent sound reasoning from his evidence which concludes in a doxastically justified belief that ~U; that is, he lacks p-independent justification to believe that ~U in the robust sense.

(c) He can form a doxastically justified belief that ~U via reasoning from p, perhaps along with other premises. That is, he can have p-dependent justification to believe ~U in the robust sense.

If someone is in this position, prior to acquiring G there is no way for anyone with his evidence to form a doxastically justified belief that ~U (that's condition (b)). However, when he acquires evidence G, he can form a justified belief that p (that follows from (a) and Anemic Conservatism). He can then form a doxastically justified belief that ~U via inference from p (that's condition (c)). But that's just Moorean reasoning.

As a concrete example, consider:

The Seasoned Interviewer Redux: As before, Sasha is an excellent and experienced police interviewer. Also as before, Sasha sincerely pleads agnosticism when a colleague asks her if a suspect who frequently touches his nose during an interview is more likely to be lying – she finds it hard to reason about such rules in the abstract and has a poor track record when doing so. Shortly thereafter, Sasha interviews a suspect who frequently touches his nose. Sasha forms a justified belief that the suspect his lying and this belief is based, in large part, on his frequent nose-touching. After the interview, Sasha remembers her colleague's question. She realizes that nose-touching seemed to her a clear sign that the subject was lying. She concludes that frequent nose-touching is a sign of dishonesty after all.

Sasha has engaged in some Moorean reasoning here. According to anemic conservatism, Sasha's reasoning is kosher. Here is why:

> (a') Prior to conducting her interview with the nose-touching suspect, Sasha has adequate evidence to believe that *frequent nose-touching is a sign of dishonesty* so according to anemic conservatism there is no obstacle to Sasha acquiring a justified belief that the suspect is lying based on his frequent nose-touching.

> (b') Prior to conducting her interview, there is no sound, undefeated bit of reasoning from Sasha's evidence to the conclusion *frequent nose-touching is a sign of dishonesty* – all such routes are defeated by Sasha's evidence that she is no good at reasoning about such rules in the abstract.

> (c') Sasha can form a doxastically justified belief in *frequent nose-touching is a sign of dishonesty* via reasoning from *the suspect is lying* and the knowledge that her belief that the suspect is lying is based upon his frequent nose-touching.

So, prior to conducting the interview, Sasha is rationally agnostic about *frequent nose-touching is a sign of dishonesty*. But, once she does conduct the interview, she can form a justified belief that the suspect is lying, and then a justified belief in *frequent nose-touching is a sign of dishonesty*. There is nothing wrong with this Moorean pattern of reasoning – at least, not according to anemic conservatism. The reason why this Moorean reasoning is kosher is that impediments to reasoning one's way to a justified belief in a proposition can be *fine-grained*: they block certain routes rather than others. Here, Sasha has evidence that she's no good at formulating rules for evaluating testimony in the abstract, but no evidence that she can't derive such rules from concrete cases. That's why she can't just reason her way to *frequent nose-touching is a sign of dishonesty* from reflection alone, but *can* do so by reasoning from her conclusion about a particular suspect. Moorean reasoning offers a detour around the impediment.

Anemic conservatism sometimes licenses Moorean forms of reasoning in humdrum cases, but that's no strike against it, since Sasha's reasoning is unobjectionable. What matters, from the perspective of the Same Boat Hypothesis, is whether anemic conservatism licenses Moorean reasoning in response to global skeptical hypotheses: that is, whether it is consistent with Moore 1-3 and Moorean Induction 1-3 being good reasoning. That question is tricky. It turns on why Mooreans think we can't form justified beliefs in not-BIV and not-NUW via apriori reflection alone. A plausible account of why not is that we should not trust ourselves to

so reason – that we have compelling grounds to doubt our ability to figure out what are surroundings are like, or what the future will be like, by apriori reflection alone. A little more carefully, we ought to believe:

Rational limitation thesis – surroundings: Humans can typically only form justified contingent beliefs about their surroundings via reasoning that, at some point, depends upon perception or perceptual beliefs.[34]

Rational limitation thesis – future: Humans can typically only form justified contingent beliefs about the future via reasoning that, at some point, depends upon induction.[35]

Where, once again, I mean *rational* rather than *psychological* dependence. If those theses are correct, there is an impediment to arriving at a justified belief in not-BIV or not-NUW via apriori reflection alone but no impediment to doing so via inference from perceptual or inductive beliefs. That's just the sort of situation wherein anemic rationalism licenses Moorean reasoning.

Why believe in the rational limitation theses? It's worth noting that something like the rational limitation theses is implicit in the premises Nasty-1 and Nasty Induction-1: those premises depend, for their plausibility, on the thought that the only ways to learn about one's surroundings and the future depend upon perception and induction. Anyone at all worried by those arguments must find the rational limitation theses plausible. Of course, that's not an argument but it means that anyone worried by the skeptical problem that motivates the (apparent) choice between Mooreanism and rationalism must find the rational limitation theses at least initially compelling. The only argument I can offer in their favor is a boring and empirical argument by elimination: I don't see other routes that we use to form justified beliefs about contingent claims about our surroundings or the future. When I observe people forming justified beliefs about their surroundings or their future, they always do so in a way that depends upon

---

[34] A possible type of counterexample includes claims such as *there are not exactly 123456789 planets in the universe*: this claim is contingent and plausibly apriori. I am inclined to say that if such beliefs are justified apriori it is on the basis of some very weak version of the principle of indifference. But thinking that we have apriori justification to believe *there are not exactly 123456789 planets in the universe* on the basis of some version of the principle of indifference does not give reason to believe that we can arrive at not-BIV through armchair reasoning; in fact, the principle of indifference (in its stronger guises) weighs against not-BIV being apriori.

[35] A complication: sometimes we can form justified belief about the future on the basis of testimony. This is no great objection to the thesis in question since the source of the testimony, to be trustworthy, must have conducted an inductive inference (or received testimony from someone who has).

perception or induction. There may be a deeper explanation of why the rational limitation theses are true or an apriori argument for them, but I don't have either one; even so, I find the empirical case compelling.

Of course, the rational limitation theses alone do not establish that Moorean responses to skepticism succeed; it must also be true that we have justification in the anemic sense to believe not-BIV and not-NUW prior to engaging in Moorean reasoning. Put another way, the rational limitation theses (and similar impediments to paths of reasoning) must be the *only* reason why there is no undefeated deliberative route from our evidence to not-BIV and not-NUW. Skeptics may balk at this and insist that, in addition, we just lack sufficient evidence for not-BIV or not-NUW. But, of course, rationalists are in no position to take that line of objection against Mooreans. And this makes rationalists ill positioned, if we ought to believe in the rational limitation theses, to say what's wrong with Moorean responses to skepticism.

3.7   CONCLUSION: WHAT KIND OF THESIS IS RATIONALISM?

Let's take stock: we have so far surveyed three possible grounds for condemning Moorean reasoning, and found that rationalists cannot adopt any of them. Each of these diagnoses thus doubly fails: it fails to offer a plausible criticism of Moorean reasoning, and in addition fails to offer (non-*ad hoc*) grounds for rejecting the Same Boat Hypothesis.

What follows from this investigation? First, we ought to be suspicious of our intuitions—intuitions that I share—that Moorean reasoning seems fishy. Skeptical arguments have a powerful hold on us: even students new to philosophy feel the force of evil demon-style skeptical arguments, and almost nobody initially finds either rationalism or Mooreanism remotely compelling. Yet it turns out to be a difficult affair to say just how these evil demon-style skeptical arguments are supposed to work, and even more difficult to articulate the ways in which Moorean replies seem viciously circular. We've managed to come up with three possible ways of spelling out this circularity, each of which suffers from a subtle problem. All of this

complexity leads me to doubt that we should trust our guts here. Perhaps we intuitively latch on to something like the insensitivity diagnosis or the dialectical inefficacy diagnosis (conservatism is probably too sophisticated to describe anyone's gut reaction) and just don't have sufficiently nuanced intuitions to pick up on the problems with these approaches. In any case, even people who want to stick with their guts and count Moorean reasoning as no good still owe us a tenable story about what's wrong with it.

Second, we have seen that there is little role for rationalism to play as an anti-skeptical alternative to Mooreanism: if rationalist responses to skepticism are kosher, they are also unnecessary because Moorean responses are kosher as well. The same boat thesis is true, so rationalism has nothing to offer as an anti-skeptical rival to Mooreanism.

One way to interpret this last result is that there is no reason to be a rationalist. I doubt, though, that that's the right interpretation. A better interpretation is that it is a mistake to think of Mooreanism and rationalism as rivals. Mooreanism is a view about *good reasoning*: it says that arguments such as Moore 1-3 and Moorean Induction 1-3 are ways to acquire doxastically justified beliefs in their conclusions. We can also interpret rationalism as a theory about good reasoning: the view that you can form justified beliefs in not-BIV and not-NUW through the simplest possible process of reasoning: you can just conclude them, and thereby be doxastically justified. We've seen that any account of what's wrong with Moorean reasoning will condemn that sort of rationalist response to a skeptic armed with VMN. However, rationalism need not be a theory of good reasoning at all. Another interpretation of rationalism is as a view about *evidential support*. Bodies of evidence can weigh for or against propositions. It's tempting to think that an empty body of evidence—one containing no evidence—must, for any deeply contingent proposition p, give just the same degree of support to p as not-p. Rationalism, if it's a thesis about evidential support, is just the denial of this claim. Rationalists think empty bodies of evidence aren't always impartial between p and not-p.

If rationalism is a view about evidential support and Mooreanism is a view about good reasoning, the two aren't obviously rivals. This is good news for rationalists, since we've seen that rationalism as a thesis about good reasoning has no important anti-skeptical role to play. It's also good for Mooreans, since there are some compelling arguments in favor of the rationalist view of evidential support and the Moorean view of reasoning does not diminish the strength of

these arguments.[36]   The conclusion of my argument isn't that rationalism is false, but rather that—when it comes to responding to skeptical arguments—it has a role to play solely as a view about evidential support; as a result, even rationalists need a story about how to reason our way to the falsity of skeptical scenarios. Failure to discredit the Same Boat Hypothesis means that story will be—in key ways—Moorean.

---

[36] See Hawthorne (2002), White (2006), and Pryor (ms). Neither Hawthorne nor White make the distinction between theories of evidential support and theories of good reasoning; however each argument weighs only in favor of rationalism understood as a view about evidential support. White takes his argument to discredit Mooreanism, but only because he assumes that, whatever Mooreanism is, it must be inconsistent with the rationalist view of evidential support – just the view of the logical space that I'm inveighing against.

4.0   SKEPTICISM AND THE ETIOLOGY OF NORMATIVE JUDGMENTS

Each of us tries to conform to norms: practical norms, which govern our actions, and epistemic norms, which govern our beliefs. A little reflection, though, can raise puzzles about whether any of us ought to believe that her norms, or her beliefs about norms, are *correct*: i.e. whether her beliefs about normative matters are true, and whether her epistemic and practical behavior corresponds with what she in fact has most reason to do. One source of skepticism is persistent disagreement: if you and I embrace different norms, why should I think that I am correct and you are wrong, rather than vice versa? That is a difficult question to answer, but even those who despair of defeating that skeptical worry can take comfort in the limited range of disagreement: humans embrace only a tiny fraction of the range of possible norms. Here I will focus on a second, more troubling kind of skepticism: rather than a worry about whether my norms rather than yours are correct, it is a worry about whether any of our normative judgments are better than blind stabs in the dark. According to this second worry, the near consensus among humans about the correct norms is no reason to take comfort; it is only a sign that we all err in similar ways.

What is the source of this second type of skepticism? The worry rests on the observation that there is a causal story about why each of us follows some particular set of norms. Consider, first, the set of causal factors that led me to embrace my norms: roughly, biological evolution, which is a mix of natural selection and purely random chance.[37] There are all sorts of ways in which selection pressures would favor some norms over others: norms that helped our early ancestors find food, build shelters, avoid predators, show off to potential mates, and so on would

---

[37] Of course, cultural influences matter as well. However, our cultures are made up of humans, each of whom is a product of natural selection. So while cultural norms are an important proximal cause of my normative judgments, evolution is the most important distal cause. So I will focus on evolution here, though the simplification will not matter very much.

contribute to survival and reproduction. Other norms would not. The skeptical puzzle arises because the *correctness* of the correct norms does not to add to their adaptive value. Our ancestors were not, simply in virtue of having justified beliefs, or doing what they had most practical reason to do, in a better position to survive and reproduce. Of course, justified beliefs or practically correct actions might also tend to have other features which did provide an evolutionary advantage, so evolution might thus have steered us towards using correct methods as a kind of fortuitous side-effect. In any case, since following correct norms as such provided no adaptive advantage, it would be an astonishing piece of cosmic good luck if evolution resulted in our following correct norms. This seems like a compelling reason to doubt that I've gotten it right – to believe that, whatever the merits of my norms, they are not correct ones.

Though this skeptical argument is intriguing, it may be dead in the water. A simple-minded reply is that if we add a little more detail to the skeptic's puzzle it loses its force. The correct epistemic norms include induction, deduction, and believing what you perceive. And it is no mystery why we evolved to favor those norms rather than counter-induction, wishful thinking, and believing what you imagine. Our norms are, at least roughly, the reliable ones. Correct practical norms include refraining from murder, keeping promises, and caring about one's own future well-being. It's no great mystery why we evolved to favor those norms rather than ones that encourage murder, untrustworthiness, and wanton disregard for tomorrow. So, our landing on roughly the right norms doesn't look so fortuitous after all.

Of course, skeptics will not be satisfied with this reply: it has a whiff of intolerable question-begging about it. The problem, roughly speaking, is that simple-minded anti-skepticism invokes normative judgments—murder is wrong, induction is rational, and so on—and there are *the very judgments that skeptics call into question*. In some sense, then, this reply begs the question against skeptics. However, it is not obvious whether skeptics can turn this worry into a damning charge against naïve anti-skepticism.

My aim is to assess the force of skepticism based in thinking about the causal origins of our normative judgments. This will require us to examine—in some detail—the ground rules for argument between skeptics and non-skeptics. In particular, we will need to get straight on the merits of the sort of naïve anti-skepticism sketched above; doing so will require investigation of the delicate question of which argumentative strategies are viciously circular, as opposed to acceptably circular. Along the way a few general lessons will emerge about which types of

skeptical argument can successfully deflect replies that are, in some sense, circular and which cannot. The upshot of my discussion will be that there had better be an explanatory relation between the normative facts, on the one hand, and our normative judgments on the other. This demand rules out non-naturalistic normative realism as well as particularist views wherein normative facts are "shapeless" at the non-normative level: each precludes the right sort of integration of normative facts into our picture of the explanatory order. Part of what is exciting about this argument—aside from its upshot about the nature of the normative—is that it presents a challenge even for "Moorean" views that are extremely sanguine about circular responses to skepticism.

Here's what's ahead: in part 1 I will describe an appealing view of the meta-normative that invites a skeptical worry. In part 2 I will present the skeptical worry more carefully; this presentation will closely resemble some skeptical arguments endorsed in recent literature. In part 3, I will argue that the standard presentation of the worry fails: anti-skeptics have a simple but devastating reply. After showing problems with the standard presentation of the argument, I will introduce some epistemic machinery in part 4 and use that machinery to present a better version of the skeptical argument in parts 5 and 6. Part 7 will specify which views of the normative can avoid the skeptical challenge. Part 8 will conclude by briefly—and inconclusively—discussing grounds for doubting the argumentative strategy of this paper.

4.1   WHO IS THE TARGET? (PART I)

The present skeptical worry stemming from the etiology of our normative judgments concerns the relations between our *normative faculties* on the one hand—i.e. whatever it is that causes us to produce the normative judgments that we do—and the normative facts on the other. I will begin describing the problem by focusing on normative views that share two assumptions. The first is **factualism**: the view that there are true normative propositions. It is clear enough that no worry about whether our normative beliefs coincide with the facts can be hung around the necks

of non-factualists. Factualism in this sense doesn't require any sort of spooky normative realm; it simply requires that sentences with normative content can be true in a minimal, disquotational sense. From now on, I will assume that factualism, in this minimal sense, is true.

The second assumption is that there is no explanatory connection between the normative facts and the output of our normative faculties. By this I mean three things:

(1) Our normative judgments do not explain why the normative facts are what they are.

(2) We do not have the particular normative faculties that we do *because of* the normative facts.

(3) There is no third factor that explains both why we have the normative faculties that we do *and* why the normative facts are what they are.

Why believe each of these claims? (1) says that our tendencies to make certain judgments do not explain why the normative facts are what they are; roughly speaking, it says that the normative facts would be the same even if we had different normative faculties. **Constructivists** may reject (1), but everyone else accepts it.

(2) says that the normative facts do not explain why anyone came to have some particular set of normative faculties. For instance, it says that I do not have an innate tendency to find murder wrong because murder is, in fact, wrong. Why believe in this assumption? The only satisfactory type of explanation of why we have the normative faculties that we do will be a *causal* explanation. Non-causal explanations can sometimes be appropriate, but not here – there is no good *rational* explanation of why we are born with certain faculties, and there doesn't seem to be any good *constitutive* explanation of why we have some normative faculties rather than others.[38] So fans of **naturalism**—the view that normative facts are identical to natural facts—

---

[38] A possible complication: can someone argue that the normative facts do play a constitutive role in determining the contents of our thoughts, and thereby show that the normative facts constitutively explain why we have the faculties that we do? Whether the normative facts play a role in determining the contents of our thoughts—that is, whether the fact that murder is wrong determines the content of what I express when I think "murder is wrong"—falls well outside of the scope of this paper (for a defense of such a position, see Wedgwood (2007)). Even if normative facts do play such a role, however, there is limited room for a constitutive explanation of why we have the faculties that we do: on such a view, normative facts play no role in explaining why we tend to think "murder is wrong" rather

can reject (2), but fans of other meta-normative views cannot. Even for naturalists, though, rejecting (2) isn't trivial: showing that normative facts are in the right ontological category to explain why we have the normative faculties that we do isn't the same as showing that they actually explain it.

Finally, (3) denies the existence of a third factor that explains both the normative facts and why we have particular normative faculties. That is, it says that there is no x such that x explains why murder is wrong, and x explains why we have the normative faculties that we do. The defense of (3) is simple: there are no obvious candidates for x.

I will call the conjunction of (1)-(3) *Explanatory Independence*, for short. Though the above considerations are not a watertight argument for Explanatory Independence, I hope they are enough to render it plausible for now (we'll re-examine the possibility of rejecting Explanatory Independence later on). Many appealing forms of **non-naturalistic** realism—views of the sort defended by Dworkin, Nagel, Parfit, Raz, and Scanlon[39] wherein normative claims are true (in the minimal, disquotational sense of "true") but normative facts are not identical to natural facts and do not depend on what we happen to believe—are committed to Explanatory Independence. In spite of its considerable plausibility, though, I will argue that Explanatory Independence has intolerable skeptical consequences. As a result, we ought to reject it. Before we can see the problem with Explanatory Independence, though, we'll need to set out our skeptical worry with a great deal more care.

4.2   A WORRY ABOUT COINCIDENCES

If normative skepticism is false we may believe that, most of the time, our normative judgments and the normative facts coincide. Let us suppose, for now, that Explanatory Independence is

---

than "murder is fine". So on such a view normative facts may play a role in determining why the output of our normative faculties count as expressing certain content, but that's not what we're looking for: an explanation of why we have normative faculties that tend to produce certain outputs.

[39] See, among other places, Dworkin (1996), Nagel (1986), Parfit (ms), Raz (1999), and Scanlon (2000).

true. On the face of it, the coincidence between our normative judgments and the normative facts is puzzling because the two coinciding elements—our normative judgments and the normative facts—are determined by totally independent factors. On the one hand, we know, in rough outline, why we humans have the normative beliefs that we do: our faculties of normative reasoning are the products of biological evolution, itself a result of natural selection and pure chance. Natural selection tends, over time, to produce traits that are conducive to the survival and reproduction of a species and to eliminate traits that are not. Thus, our proclivities to make certain normative judgments must be ones that, on the whole, contributed to the survival and reproductive success of our ancestors.[40] Our ancestors evolved normative faculties in response to certain properties of actions and beliefs; our judgments are certainly not random. The problem is that our ancestors could not have evolved particular normative faculties *because* those faculties produce largely true judgments: Explanatory Independence assures us as much. On the other hand, our judgments play no role in determining the normative facts: Explanatory Independence assures us of this much, too.

In general, when the two elements of some putative coincidence are explanatorily independent of one another, we ought not to believe that they coincide. For instance, suppose I decide to make a telephone directory for Bangor, Maine, by coming up with names using a Ouija board and assigning each name a number based on seven consecutive rolls of a ten-sided die. The phone numbers of people in Bangor do not affect the listings in my book, and the listings in my book do not affect the phone numbers of the people in Bangor. That the phone numbers and the listings in my book are determined independently doesn't *guarantee* that my book is inaccurate, but it certainly renders any putative coincidence between my listings and the real phone numbers puzzling. One way to spell out our skeptical worry, then, is that our normative judgments are like the homemade phonebook: it would be an astonishing coincidence if they were right.[41] We can present this worry a bit more carefully:

---

[40] Another possibility is that our normative faculties have no survival value and arose as a fluke, but that possibility won't help us to resist skepticism.

[41] A skeptical argument very much like Coincidence 1-5 is presented by Enoch (forthcoming), Schafer (forthcoming), and Street (forthcoming) and (ms), though Enoch and Schafer do not ultimately endorse the argument. Field (1991) uses an argument along roughly these lines against mathematical Platonism – see, especially, the introduction.

Joyce (2006) and Street (2006) present related worries – I'll say a bit more about Street's "Darwinian Dilemma" in section 6.

Premise: (Coincidence-1) Explanatory Independence
Premise: (Coincidence-2) If Explanatory Independence, then the alleged coincidence between our normative judgments and the normative facts is inexplicable.
So: (Coincidence-3) The alleged coincidence between our normative judgments and the normative facts is inexplicable.
Premise: (Coincidence-4) It is inappropriate to believe in an inexplicable coincidence between our normative judgments and the normative facts.

Therefore: (Coincidence-5) We ought not to believe that our normative judgments and the normative facts coincide.

Two notes about this argument. First, this style of argument does not threaten to generalize to render all coincidences between our beliefs and the facts puzzling. Consider ordinary perceptual beliefs. The features that those beliefs are about—color, shape, size, and so on—are definitely well-positioned to figure in explanations of why we have the perceptual faculties that we do. It is easy to envision, in rough outline, a causal explanation of the coincidence between the output of our perceptual faculties and the facts about material objects.[42] Second, I raised the skeptical worry behind Coincidence 1-5 by mentioning evolution by natural selection. However, Coincidence 1-5 does not mention evolution. Thinking about evolution by natural selection makes the skeptical worry vivid, but does no more. It does not matter what causal process led us to have the normative faculties that we do. All that matters is that the process does not allow the normative facts to explain the output of our faculties; that is, does not allow us to say that we believe that murder is wrong *because* murder is wrong. If we were the products of Lamarckian evolution, say, just the same skeptical problem would arise. Coincidence 1-5 is apriori.

Of course, not everybody needs to be troubled by this argument: meta-normative views that reject Explanatory Independence can reject the first premise. Fans of Explanatory Independence, though, appear to face a problem. The premise Concidence-2 is obviously correct. Coincidence-4 looks strong as well. It says that *if* the alleged coincidence between our judgments and the normative facts is inexplicable, then it is inappropriate to believe in it: trusting

---

[42] Arguments with roughly the shape of Coincidence 1-5 *may* raise skeptical doubts about other domains of belief. First, the same worry might arise over our beliefs about causally inefficacious domains such as mathematics and logic. Second, an analogous argument might make trouble for beliefs about subject matters that our distant ancestors didn't need to know about: distant stars, subatomic particles, and so on. These are big issues worthy of separate treatment, so I'll set them aside here.

those judgments is epistemically on a par with trusting the Ouija-and-dice generated Bangor phonebook.

How bad is the conclusion, Coincidence-5? Very bad. We should distinguish Coincidence-5 from the mild skeptical claim that some distorting factors affected the evolution of our normative faculties. *That* result is obvious: of course nobody should treat his judgments as a direct and infallible pipeline to the normative truth. However, if Coincidence-5 is true none of us can trust any of his normative judgments at all. The problem is not just that they might be false—it is obvious that we have no guarantee of truth—but that it would be unreasonable to trust them. The only rational reply would be agnosticism about all normative matters, and that is a radical position indeed. Coincidence 1-5 thus presents a formidable challenge. How should anti-skeptics reply?

4.3   NAÏVE ANTI-SKEPTICISM

The argument Coincidence 1-5 fails. A skeptic pushing Coincidence 1-5 tries to defeat normative knowledge by arguing that non-skeptics must believe the unbelievable: namely, that there is an enormous inexplicable coincidence between our normative judgments and the normative facts. This skeptical strategy requires the premise:

> (Coincidence-4) It is inappropriate to believe in an inexplicable coincidence between our normative judgments and the normative facts.

The problem with Coincidence-4 is that it is often appropriate to believe in inexplicable cosmic coincidences. For instance, suppose I buy one lottery ticket—the only one I've ever bought—and the number on *my* ticket matches the winning number. What are the odds! Certainly I should doubt very much that such a thing will happen: when I buy the ticket, I should not believe that the number on my ticket and the winning number will coincide. However, once the number on my ticket matches the number on TV, and I go to the lottery office and they say I've got a

winner, and reporters show up at my house, and friends start calling and asking for money, and checks from the lottery start arriving in the mail I may certainly believe in the enormous, unexplained coincidence.

Perhaps the lottery example does not seem like a grand enough coincidence: after all, someone had to win. But consider another, more puzzling coincidence: the coincidence between the actual physical constants and the physical constants that render life possible. Had just about any physical constants been even a little different—gravity a little stronger, the strong force a little weaker, and so on—the universe would have either collapsed shortly after the Big Bang or quickly spread out into cold, boring soup of isolated particles. Many, many constants had to fall within a very narrow range to render stars and planets, let alone life, possible. This coincidence is as enormous as they come, and may turn out to be inexplicable as well. But it would be crazy to become agnostic, on these grounds, about whether our universe is suitable for life.

Skeptics may not be fazed by these counterexamples. In the case of the lottery ticket or our physical laws, we have lots of evidence for the coincidence: checks coming in the mail, life all around, and so on. But, skeptics may insist, there is nothing analogous in the case of the coincidence between normative judgments and normative facts. *If* you had some analogous evidence that you were getting things right, it would be fine to believe in the coincidence; but you don't, so it isn't.[43]

There is an obvious, simpleminded reply to this objection. Evolution led us to believe that deliberate cruelty is wrong, and deliberate cruelty is indeed wrong, so look: evolution led us the right way. Evolution led us to believe that induction generates justified beliefs, and induction does generate justified beliefs, so look: evolution led us the right way again. This sort of vindication is not as cheap as it may look at first. For instance, evolution appears to have led many people to find beliefs supported by the gambler's fallacy plausible, even though nobody, after reflection, thinks that the gambler's fallacy generates justified beliefs. Similarly, evolution appears to have led most of us to care less about the suffering of people located far away, even though few of us want to endorse the principle that the suffering of people far away really does matter less. We could imagine making a pattern of normative judgments that was incapable of being tamed into a coherent picture, and someone making such judgments would not be able to

---

[43] The lottery example, and a reply along the lines sketched here, comes from Street (ms).

reasonably believe that evolution led him the right way. In actuality, however, these bits of tension seem to be relatively few and far between – it looks as if, most of the time, evolution led us to be inclined to get it right.[44]

Here skeptics will cry foul. This vindication, they argue, is circular. But here the anti-skeptic should dig in his heels. Circular how? Consider the argument:

> **Premise:** (Cruelty-1) Evolution led us to judge that deliberate cruelty is wrong
> **Premise:** (Cruelty-2) Deliberate cruelty is wrong
> **Therefore:** (Cruelty-3) Evolution led us to correct normative judgments about cruelty.

This argument is not **tautological**: its conclusion does not appear among the premises. In what, then, does its alleged circularity consist?

A possible answer is that the argument uses as a premise Cruelty-2, which normative skeptics claim we cannot reasonably believe; thus the anti-skeptical picture is based on dubious foundations. If we cannot reasonably believe Cruelty-2 and similar normative claims, we have no basis to believe in a coincidence between our judgments and the normative facts. Thus, the anti-skeptical position cannot be consistently held. However, this charge is too quick. Certainly someone who accepts Coincidence-5 should not trust his normative judgments, and someone who (for whatever reason) does not trust his normative judgments cannot engage in reasoning along the lines of Cruelty 1-3. Similarly, skepticism about individual normative claims (such as Cruelty-2) leaves one unable to respond to Coincidence 1-5, but that does not show that the rest of us must accept Coincidence-5. It *does* show that accepting one skeptical claim but not the other—i.e. being skeptical about ordinary normative judgments such as Cruelty-2 but rejecting Coincidence-5, or *vice versa*—is an untenable position. However, there is no inconsistency in the unabashed non-skeptical stance: accepting Cruelty-2 (and other ordinary normative claims) and rejecting Coincidence-5. If the skeptic's aim is to show inconsistency within the common sense position, he fails.

Skeptics might want to say something else: not that the non-skeptical position is internally inconsistent, but that invoking ordinary normative judgments such as Cruelty-2 as part of a response to Coincidence 1-5 violates some dialectical rule. Trusting ordinary normative

---

[44] Something like this reply to a similar worry is offered in Dworkin (1996).

judgments in some sense presupposes that skepticism is wrong. Invoking such judgments to argue against skeptics feels non-kosher; it's a kind of question-begging. In one sense, this response is correct: a rational skeptic need not be convinced by the naïve anti-skeptical strategy. If the game is convincing rational skeptics, the naïve response loses.

However, the inability of naïve anti-skepticism to rationally convince skeptics is no great failing, for two reasons. First, this failing does not rule out the use of the naïve reply as a reason for non-skeptics to continue to trust their normative judgments. We cannot rule out any reply to skepticism that uses the judgments that skeptics call into question: that would make victory for the skeptic all too easy. Consider a skeptic who asserts that nobody can reasonably believe anything. Any reply to this sort of skeptic is question-begging in the sense that it will use one of the very claims that skeptics deny we can reasonably believe as a premise in an argument against skepticism. If that sort of circularity is always bad, the skeptic wins; indeed, he wins without giving any argument at all. That cannot be right. The upshot is that skeptics cannot rule out-of-bounds the practice of defending non-skeptical positions by invoking judgments that skepticism calls into question. While skeptics do not have to be convinced by the naïve reply, they have not shown that non-skeptics cannot feel secure in their own views as a result of giving the naïve reply. If the game is rationally defending non-skepticism, the naïve response succeeds.

Second, rationally-convince-the-skeptic is a sucker's game. The problem is not that it is hard – though, of course, it is. Rather, the problem is that it is pointless even if it succeeds. Skeptics offer only *ad hominem* arguments: they try to show that common sense is inconsistent. But skeptics do not adopt internally consistent positions and this is why skeptics, since antiquity, have denied that they mean to defend any position at all. In the present case, a skeptic pushing Coincidence 1-5 asserts a normative claim about what it is appropriate to believe: specifically, Coincidence-4. Someone who denies that it is rational to believe any normative claims and yet believes Coincidence-4 has an incoherent position. So it would be no great feature of an anti-skeptical strategy that it could rationally convince a skeptic: that would mean that a rational skeptic would have to abandon her position on pain of inconsistency, but *even absent an anti-skeptical reply* a rational skeptic must abandon her position on pain of inconsistency. *That's* why skeptics have to deny that they are defending a position. So winning a game of rationally-convince-the-skeptic would not get anti-skeptics anything that they do not already have.

On the one hand, skeptics lose this round. Anti-skeptics have an internally consistent position: someone who trusts his normative judgments can deny Coincidence-5 and thus appropriately retain his normative judgments. A skeptic cannot rule out the naïve line of reply. On the other hand, it is easy to sympathize with the skeptic's complaint that his opponent's rather breezy and circular reply does not get to the heart of the worry. Skeptics are right about this, but the problem lies not with the anti-skeptic's reply—which, while breezy, certainly works—but with Coincidence 1-5. That argument has little skeptical force, but there is another worry in the same neighborhood that is far more powerful. Before we can state it, and see why it defeats naïve anti-skepticism, we need to get straight on evidential support and defeat.

4.4   EVIDENTIAL SUPPORT: HOW TO GET IT AND HOW TO LOSE IT

Before stating a more powerful skeptical argument we need three pieces of epistemic machinery. First, we need a theory of evidential support. Under what circumstances does some evidence E provide rational support for p? That is, what relationship must E and p have in order for learning that E to make belief that p more reasonable? One natural view is that E is evidence for p if and only if there is some explanatory relation between E and p. I will officially remain agnostic about the *if* direction of this claim, but I will embrace the *only if* direction, i.e. I will accept:

> **Explanatory Constraint:** E can be evidence for (or against) p only if there is some explanatory relation between E and p.

Where E is *explanatorily related* to p just in case *either* S learning that E makes p a better (or worse) explanation of S's total evidence than it was before he learned that E *or* S learning that E makes S's total evidence a better (or worse) explanation of p than it was before. Put another way, if learning that E does not make p a better (or worse) explanation of one's total evidence or

77

make one's total evidence a better (or worse) explanation of p, $Pr(p|E) = Pr(p)$.[45] Two clarifications about the Explanatory Constraint: first, it tells us about relations between propositions in general - it does not require that the relevant propositions must be true. We can have misleading evidence. Second, there are several ways for propositions to be explanatorily related. If p explains q then p can be evidence for q and q can be evidence for p. A fire in the fireplace explains why there is smoke coming out of the chimney: so *there is a fire in the fireplace* can be evidence for *there is smoke coming out of the chimney* and *there is smoke coming out of the chimney* can be evidence for *there is a fire in the fireplace.* In addition, if q and r have a common explanation, q can be evidence for r and r can be evidence for q. A fire in the fireplace explains heat emanating from the fireplace and smoke coming out of the chimney, so *there is smoke coming out of the chimney* can be evidence for *there is heat emanating from the fireplace*, and *there is heat emanating from the fireplace* can be evidence for *there is smoke coming out of the chimney*.

Why believe in the explanatory constraint? The answer lies in rationality of making an inference to the best explanation. Suppose that S has total evidence T, and the best explanation for all of this evidence is p, rather than some mutually exclusive alternative hypothesis q (for simplicity, we'll assume p and q are the only contenders). So, S forms a justified belief that p via inference to the best explanation.   Then, S acquires a new bit of evidence, E, which is explanatorily unrelated both to whether p and to whether q. Suppose, for *reductio*, that E is powerful evidence for q: powerful enough to make it rational to abandon belief in p and replace it with belief in q. If S ought to believe the best explanation of his evidence, this move is irrational: by the definition of explanatory relevance, learning that E does not make q a better explanation of S's total evidence than it was before (and doesn't make p a worse explanation than it was before) so p is still the best explanation of S's total evidence. Inference to the best explanation demands that S continue to believe that p. So we have a contradiction: it both is and is not possible that it is rational for S to abandon p in favor of q. Thus we ought to reject our assumption: that E can be powerful evidence for q. The same argument applies, albeit less dramatically, if we imagine that E is somewhat less powerful evidence for q and thus only affects

[45] By the shorthand *Pr*(x), I mean the evidential probability of x. A bit more carefully, the evidential probability of x is a measure of the strength of the evidence that bears in favor of x, divided by the strength of all the possible evidence that could bear in favor of x. We need this latter condition because the "evidential probability" of e.g. logical claims is 1, even for someone who has no evidence in their favor.

the degree to which S ought to be confident in p and in q. If each of us ought to accord our degrees of belief according to how well each hypothesis explains our evidence, becoming less confident in p upon learning E will be rational only if E is explanatorily related to p. Inference to the best explanation and the Explanatory Constraint go hand in hand: anybody who rejects the latter must reject the former. Since rejecting inference to the best explanation is a radical move, and the Explanatory Constraint is independently plausible, we ought to embrace both.

The Explanatory Constraint may look familiar: it resembles, to some degree, Gilbert Harman's "explanatory relevance test".[46] Certainly the two constraints share some features in common, but the Explanatory Constraint is weaker than Harman's test in three ways. First, Harman says that we have evidence for p only if p helps explain the occurrence of some observation. The Explanatory Constraint does not share this exclusive focus on relevance to explaining the occurrence of observations: E is evidence for p only if there is an explanatory relation between E and p, but E needn't be the occurrence of an observation. Second, Harman's test requires that some claim figure in the *best* or *simplest* explanation of an observation; at least, his argument requires as much.[47] The Explanatory Constraint simply requires *some* explanatory relation. Third, Harman allows only one kind of explanatory relevance: an observation is evidence for p only if p helps explain that observation. The Explanatory Constraint allows for three kinds of explanatory relevance: an observation is evidence for p if p helps explain that observation, or if the observation explains p, or if some common third factor explains both p and the observation. Together these make the Explanatory Requirement far less radical than Harman's test.

So much, then, about how to acquire evidential support; we also need an account of how to lose it. Suppose Samantha has a belief that p that is **based**—psychologically dependent—upon some evidence E; E is strong enough evidence to render Samantha's belief that p justified. A **defeater** for p is a proposition such that, were Samantha to come to reasonably believe it, she would no longer have justification to believe that p. One type of defeater is an **outweighing defeater**: a proposition that provides evidence for the falsity of p and thus renders belief in p inappropriate in spite of the reasons given by E for counting p true. An outweighing defeater leaves the rational support that E lends to p unchanged. Another type of defeater is an

---

[46] See Harman (1988).
[47] For more on this point, see Sturgeon (1988) and Quinn (1986)

underminer. Learning that an underminer is true eliminates the rational support that E lends to p. For instance, suppose I'm waiting for some dinner guests and I hear the ring of a doorbell. That evidence gives me justification to believe that my guests have arrived. Then I open the door and see that nobody is standing there: my justification to believe that my guests have arrived has been *outweighed*. Suppose instead that learn that the doorbell sound came from the television: my justification to believe that my guests have arrived has been *undermined*.[48]

We can differentiate outweighing and undermining defeaters in formal terms. If O is an outweighing defeater for p, then

$$Pr(p|O) < Pr(p)$$

since O is evidence for ~p. If one learns that O, one ought to become less confident that p. If U is an undermining defeater for the support lent by E to p, then

$$Pr(p|U \text{ and } E) = Pr(p)$$

U is not evidence for ~p; rather, it removes whatever support E lends to p. If one learns that U, he ought to be no more confident in p than he would be if he lacked evidence E; in other words, U removes whatever *prima facie* support E lends to p.[49]

Finally, we can combine our theory of evidential support with our taxonomy of evidential defeat. The theory of evidential support outlined above allows us to transform a worry about an inexplicable coincidence into a worry about undermining. So far we have established, via the Explanatory Constraint, that

$$Pr(p|E) > Pr(p) \text{ only if there is an explanatory connection between E and p}$$

---

[48] Many propositions diminish the rational support that E lends to p to some degree. I will use the term *underminer*, though, only to refer to propositions that eliminate the rational support that E lends to p entirely.
[49] For (much) more on the formal treatment of defeaters, see Kotzen (ms).

In addition, we have established that U undermines the evidential force of E if

$$Pr(\text{p}|U \text{ and } E) = Pr(\text{p})$$

These claims, together, entail if I learn that there is no explanatory connection between E and p, the support lent by E to p is undermined. Since I ought to respond to underminers by dismissing any *prima facie* support lent by E to p, learning about the lack of explanatory connection between p and E entails that I ought to dismiss any *prima facie* support lent by E to p. In the next section we will examine the skeptical implications of this result.

4.5   UNDERMINING AND SKEPTICISM

Earlier we examined a failed skeptical argument, the crux of which was a worry about an inexplicable cosmic coincidence between our judgments and the normative facts. We can now state a related argument, the crux of which is that the output of our normative faculties provides no evidential support to any normative proposition. Why not? Because, according to the views outlined in the previous section, any such *prima facie* evidential support is undermined.

I'll show a version of this skeptical worry that makes trouble for a particular view of normative epistemology, wherein justified normative beliefs are based on evidence from the output of our normative faculties (in the next section, I'll extend the skeptical argument so that it applies to alternative epistemologies of the normative). There is more than one form for such basing to take. One form is that normative faculties produce something non-doxastic—normative intuitions, say—upon which normative beliefs are based. Another form is for normative faculties to spit out basic normative beliefs, which serve as the evidence for other non-basic beliefs; those normative beliefs are either self-basing, or else mutually based on one

another. The difference between these two models—and their multiple variants—need not concern us here. What matters is that the output of faculties, a kind of psychological fact, is evidence for normative beliefs. We now have the tools to challenge this view:

> **Premise:** (Undermining-1) Explanatory Independence
> **Premise:** (Undermining-2) If Explanatory Independence, then there is no explanatory connection between the output of anyone's normative faculties and the normative facts
> **So:** (Undermining-3) There is no explanatory connection between the output of anyone's normative faculties and the normative facts
> **Premise:** (Undermining-4) If there is no explanatory connection between the output of anyone's normative faculties and the normative facts, then any *prima facie* evidential support lent by the output of anyone's normative faculties to any normative claims is undermined.
> _____
> **Therefore:** (Undermining-5) Any *prima facie* evidential support lent by the output of anyone's normative faculties to any normative claims is undermined.

The conclusion of this argument is not that all of out normative judgments are false, nor even that we should not believe them. Instead, the conclusion is something weaker: the outputs of people's normative faculties do not provide *any* evidence for *any* normative conclusion. For instance, that you believe murder is wrong, or have the intuition that murder is wrong, or are inclined to judge that murder is wrong is no evidence that murder is wrong. The same goes for the output of other people's normative faculties: that Socrates believes murder is wrong is no evidence that murder is wrong.

Undermining 1-5 has the same rough shape as Coincidence 1-5: does the same sort of objection plague it? A similar kind of naïve anti-skepticism may look promising: I know that my normative faculties lead me to believe that murder is wrong, and murder is indeed wrong; I know that my normative faculties lead me to believe that induction is rational, and induction is indeed rational, and so on. Taking all of this evidence together: the track record points pretty clearly towards our having reliable normative faculties. If we've got good reason to believe that our normative faculties are reliable then their output is evidence for normative claims. Thus it is not at all obvious that the evidential force of the output of our faculties is undermined.

One worry about this line of response is that it involves an odious form of **bootstrapping**: using some the output of some faculty or device as part of an inductive argument for the reliability of that faculty or device. For instance, suppose I trust the output of some thermometer, but someone else challenges its reliability. It will do no good to respond by saying

"yesterday the thermometer said it was 48º and indeed it was 48º (how do I know? I checked the thermometer!); two days ago the thermometer said it was 61º and indeed it was 61º (how do I know? I checked the thermometer!)…and so on, so the thermometer must be quite reliable". However, the present case differs from obviously ineffectual forms of bootstrapping in at least two ways. First, while bootstrapping is no way to come to believe in the reliability of a particular device, say, it is not clear that bootstrapping is inappropriate when it involves some entire domain of beliefs. Skepticism is just too easy to come by if I am not allowed to use the output of my normative faculties in an argument for the reliability of my normative faculties, or the output of my material-object-belief faculties in an argument for the reliability of my material object beliefs, and so on. Second, in the thermometer case there is no possibility of getting a different result. I know in advance—before ever taking a reading from the thermometer—that the test of comparing the thermometer's output with the temperature (as determined by the thermometer's output) will end up yielding a perfect track record for the thermometer. The test thus adds nothing. However, things could have turned out differently with our normative faculties: for instance, each of us might have been inclined to endorse a bunch of inconsistent normative judgments, or endorse quite different normative judgments at different times, or find other reasons for suspicion contained within our normative judgments. In practice there is a bit of inconsistency—nearly everyone has normative intuitions that, upon reflection, she is unwilling to endorse—but on the whole each person's normative judgments look internally consistent enough to be largely, if imperfectly, trustworthy. The question of what, precisely, makes a bit of reasoning an odious form of bootstrapping is delicate, but it is far from clear whether charges of bootstrapping will stick against the simple anti-skeptical move.[50, 51]

There is a more damning worry about the present anti-skeptical strategy: it relies on the very normative judgments that, according to Undermining-5, we should not trust. This feature of the anti-skeptical reply is familiar. Once again, the question is whether the anti-skeptic's use of a normative judgment as part of an objection to a skeptical argument constitutes a *vicious*

---

[50] For more on bootstrapping, see Weisberg (forthcoming) and White (2006).

[51] Was the naïve reply to Coincidence 1-5 an instance of bootstrapping? I think not. One difference between that reply and canonical cases of bootstrapping is that the naïve reply to Coincidence 1-5 does not draw any claims about reliability from the observation that his normative beliefs have largely been true. Weisberg (forthcoming) argues that the inference from a claim about a track record to a claim about reliability is *precisely* where bootstrapping goes wrong: if he is right, the reply to Coincidence 1-5 is not fine even if the reply to Undermining 1-5 is not.

circularity. We are assuming, for now, that the relevant normative judgments—murder is wrong, induction is rational, and so on—are based on the output of normative faculties (we will relax this assumption in the next section). The question of whether the anti-skeptical move is viciously circular turns out to be: is it acceptable for an anti-skeptic to use a normative claim, p, as part of an argument for the falsity of an underminer, U, which undermines his evidential support for p? It turns out that there is a substantial literature on this very question. Some philosophers, called **Mooreans**, believe that you can use a belief p based on evidence E as part of a deliberative route that concludes with a newly justified belief that not-U. Mooreans are so named because G. E. Moore's "proof" of the external world has just that structure: he uses the belief that he has hands, based on perceptual experience, as part of a deliberative route that (according to Moore) concludes in a newly justified belief in the falsity of an underminer for the support lent by perceptual experience to the belief that he has hands (i.e. the falsity of the claim *there is no external world*). Mooreanism is quite controversial: many, many philosophers believe that Moore's proof is viciously circular.

Fortunately for us, but unfortunately for fans of the present anti-skeptical strategy, we do not need to settle the dispute between Mooreans and their foes here. The dispute about Mooreanism is a dispute about what happens when somebody is *rationally agnostic* about an underminer. Mooreans think that you can use a belief that p based on E as part of a deliberative route that concludes in a justified belief that U—about which you were previously rationally agnostic—is false. [52] Anti-Mooreans disagree: they believe that one needs to have p-independent justification to believe that not-U in order to conclude that p on the basis of E. But all parties to this dispute agree that anyone who has a justified belief in U cannot conclude that p on the basis of E, and certainly cannot use a belief that p as part of an argument for the falsity of U. For instance, suppose I have good reason to believe that I am a brain in a vat: perhaps the words YOU ARE A BRAIN IN A VAT keep scrolling along the bottom of my visual field. In *that* case, even Mooreans do not think I can use Moore's "proof" to determine that I am not a brain in a vat; even Mooreans believe that people who have good reason to believe that they are brains in vats should not trust their perceptual judgments. If there is a genuine reason to believe in some underminer for the support lent by the output of our normative faculties to normative

---

[52] For more on the notion that Mooreanism is about the rational response to agnosticism about underminers, see Pryor (2004).

propositions, nobody can use normative propositions as part of an argument for the falsity of that underminer. The naïve anti-skeptical move doesn't work here.

We can generalize this point and distinguish two varieties of skeptical argument that focus on undermining, also called "Academic" skepticism. The first, more familiar variety of argument—including typical evil demon/brain-in-a-vat/dreaming skepticism—involves three moves. First, it raises some undermining skeptical hypothesis, U. Second, it claims that if someone is agnostic about U, and U undermines the support lent by E to p, he ought to be agnostic (at best) about p. Third, it points out that the support for the claims needed to argue for the falsity of U—for instance, perceptual beliefs—is undermined by U and that therefore we ought to be agnostic about U (and thus agnostic, at best, about p). Whether or not this strategy succeeds is a fascinating question; though few philosophers are convinced by this sort of argument—skeptics are rare—there is a remarkable lack of consensus about just where it goes wrong. A second, less common skeptical strategy provides an *argument* for U using premises that non-skeptics accept. It then claims that if someone ought rationally to believe that U, and U undermines the support lent to E by p, he ought to be agnostic (at best) about p. Finally, it points out that the support for the claims needed to argue for the falsity of U is undermined by U. I will call this second type of argument **motivated Academic skepticism**, since it involves giving affirmative motivation for U. I will call the former, more common type, **unmotivated Academic skepticism.**[53] The way to respond to motivated Academic skepticism is to challenge the motivation for U *without relying on the claims whose support is undermined*. The bad news is that motivated Academic skepticism is thus very hard to defeat. The good news is motivated Academic skeptical hypotheses are hard to come by: nobody has a colorable argument that I am a brain-in-a-vat, or whatever. However, if Explanatory Independence is correct then Undermining 1-5 is a piece of motivated Academic skepticism.

This account of the difference between motivated and unmotivated Academic skepticism explains why Undermining 1-5 is more successful than traditional external world skepticism, but does not explain why Undermining 1-5 is more successful than Coincidence 1-5. In each case the skeptical hypothesis is motivated. The difference is, roughly, that Coincidence 1-5 raises

---

[53] Skeptical arguments are not motivated or unmotivated *tout court*. Rather, they are motivated or unmotivated relative to a particular target: i.e. a skeptical argument is motivated if the target of the argument accepts the premises used to argue for the undermining hypothesis.

skeptical doubt too late. The strategy of a skeptic armed with Coincidence 1-5 is to point to all of our normative judgments and raise the worry that the (alleged) coincidence between those judgments and the normative facts is mysterious and therefore worrying. But he is not obviously right, since those normative beliefs contain the tools to demystify the coincidence. The skeptical argument tries to point out a tension in our ordinary position, but fails to do so: while it *may* establish that there is a consistent skeptical alternative, it cannot show that our ordinary position is incoherent. In contrast, a skeptic armed with Undermining 1-5 can point to each of our normative beliefs individually and argue that its (alleged) basis does not, in fact, provide it with any evidential support. She claims there is no way we could have gotten justified normative beliefs *in the first place*: those beliefs are based on normative beliefs, intuitions, or inclinations to judge, but the support of that basis for those beliefs is undermined.

Undermining 1-5 thus has two powerful features. First, it is an Academic skeptical argument: the skeptical hypothesis Undermining-5 undermines the rational support of all of the propositions that anti-skeptics might use to defend their position. Second, it is motivated: it does not just present an Academic skeptical hypothesis, but provides an argument for it based on premises that non-skeptics accept. This combination of features generates a compelling skeptical threat indeed. Of course, all of this need only worry someone who claims that justification for our normative beliefs depends upon evidence in the form of the output of our normative faculties. Can an anti-skeptic evade the force of Undermining 1-5 by adopting another view of the epistemology of the normative, wherein the output of our normative faculties need not be evidence for normative claims?

4.6   DO OTHER VIEWS OF NORMATIVE EPISTEMOLOGY FARE ANY BETTER?

What other forms can an epistemology of the normative take? All views must fall into (at least) one of three categories:

(i)   justified belief in normative claims is not based on any evidence whatsoever

(ii)   justified belief in normative claims is based on non-normative evidence

(iii)   justified belief in normative claims is based on evidence in the form of the normative facts themselves[54]

These views are not mutually exclusive—someone can think that justification for normative beliefs comes from more than one source—but they are collectively exhaustive. Support for normative claims must bottom out in normative evidence, non-normative evidence, or no evidence at all. Undermining 1-5 does not, on its own, make trouble for views of types (i)-(iii). However, if Explanatory Independence is true then these views don't dodge the underlying skeptical worry – all they do is shift the problem.

The trouble for views of types (i)-(iii) lies in the conditional

Trust Requirement: If [someone can form a justified belief that p on the basis of process R] then [the output of R is evidence that p]

The motivation for the Trust Requirement is simple: either it's reasonable to believe that some process R concluding with my believing that p indicates the truth of p or it isn't. If it is, then the output of R is evidence that p. If it isn't, I cannot form a justified belief that p on the basis of R. Either way the Trust Requirement holds. This is easiest to see with examples: if I can form a justified belief that it is 3:00 on the basis of consulting a clock, then the clock saying that it is 3:00 is evidence that it is 3:00, if I can form a justified belief in a logical claim p on the basis of Russell's testimony then Russell asserting some logical claim p is evidence that p, and so on. Note that the Trust Requirement applies to apriori domains: for instance, logic is apriori if anything is, but the assertions of trustworthy logicians such as Russell can still be evidence for the truth of logical claims. To say that logic is apriori is to say that there is a way of arriving at logical truths that does not require basing on empirical evidence; it is not to say that there can be no evidence for logical claims.

---

[54] This taxonomy follows Setiya (forthcoming).

The Trust Requirement, combined with the results of the previous section, makes trouble for views of types (i)-(iii). First, take (i): the **rationalist** view that justified belief in normative claims is not based on any evidence whatsoever. If rationalism is correct, then if I am reasoning well I will reach correct normative conclusions (even absent evidence for those conclusions). However, error is possible on such views: not everyone reasons well. Skeptics can make trouble for rationalism if they can provide grounds to doubt that one is reasoning well. And it turns out that they can. We know that there is no explanatory connection between my using apriori reasoning to reach the normative conclusion that p, on the one hand, and p, on the other. According to the Explanatory Constraint, then, Pr(p|I use apriori reasoning to reach the normative conclusion that p) = Pr(p). Thus, any *prima facie* evidential support lent by *I use apriori reasoning to reach the normative conclusion that p* to p is undermined. The conclusions of my apriori reasoning do not provide evidence for any normative claim, and neither do any else's. According to the Trust Requirement, *if* I could form justified beliefs in normative claims on the basis of apriori reasoning, *then* the conclusions of my apriori reasoning would be evidence for or against those claims. By *modus tollens*, I cannot form justified beliefs in normative claims on the basis of apriori reasoning.

Can rationalists resist this argument by saying that justification for normative claims is had by *default*, and thus isn't based on apriori reasoning or any other psychological process? No. The problem isn't that there can't be default justification; rather, the problem is that normative beliefs don't fit the bill. If normative beliefs were default, tampering with psychological processes would leave the justification untouched. Yet if I found out that all of my current moral beliefs are the result of brainwashing, say, I ought to lower my confidence in them. This shows that if I shouldn't trust my apriori reasoning, justification for apriori beliefs is threatened.

Type (ii) views, which hold that justified belief in normative claims is based on evidence in the form of non-normative propositions, face the same sort of problem as view of type (i). On type (ii) views, too, error must be possible: someone can draw false normative conclusions from non-normative evidence. Someone's normative conclusions need only be true if he both has the right non-normative evidence *and* reasons well from that evidence to normative conclusions. Here, too, skeptics can provide grounds to doubt one's normative conclusions if they can provide grounds to doubt that one is reasoning well. And, once again, they can. There is no explanatory connection between *I draw the normative conclusion that p from non-normative evidence E* and

p, therefore the former is no evidence for the latter. We can use the Trust Requirement to perform another *modus tollens* and conclude that I cannot form justified beliefs in normative conclusions on the basis of inference from non-normative evidence.

Finally, type (iii) views—wherein the evidence for normative claims is the normative facts themselves—fare no better. Of course, there may be explanatory relations between normative fact and normative fact. If I thought that I was reliably latching on to the normative facts, then, there would be an explanatory connection between the contents of my judgments and the normative facts. However, if we know there is no explanatory connection between my judgments and the normative facts, *that I judge* some normative proposition to be among my evidence provides no support at all for the claim that it *is* among my evidence. While it is true on this view that our evidence does provide support for normative claims, it is not true on this view that we should trust ourselves to figure out what evidence we have. A little more carefully, on this view Pr(p|I count normative proposition p as among my evidence) = Pr(p). Thus, given the Trust Requirement, I cannot form justified beliefs that some normative proposition is among my evidence.

Is there room for a naïve reply to these skeptical attacks? We saw in the previous section that there is no room to deny that the evidential force of someone reaching a normative conclusion—either via apriori reasoning or via reasoning from non-normative premises—is undermined; that would constitute a Moorean response to motivated Academic skepticism. Perhaps, though, there is room to deny the Trust Requirement. Consider our earlier naïve anti-skeptical argument:

> **Premise:** (Cruelty-1) Evolution led us to judge that deliberate cruelty is wrong
> **Premise:** (Cruelty-2) Deliberate cruelty is wrong
> **Therefore:** (Cruelty-3) Evolution led us to correct normative judgments about cruelty.

After conducting a bunch of reasoning along the lines of Cruelty 1-3 we might convince ourselves that our normative beliefs are trustworthy – after all, we keep getting things right! Can this generate a standoff with a skeptic armed with the observation that our normative judgments lack evidential force and the Trust Requirement?

No. We have seen that there is no room to claim that our normative judgments are evidence for normative propositions. That claim, combined with the Trust Requirement, leads to

the conclusion that it is not reasonable to trust my normative judgments, including Cruelty-2. If a non-skeptic accepts the Trust Requirement, the arguments against views of types (i)-(iii) are instances of *motivated* Academic skepticism. Responding to such arguments by relying on the very claims whose support they undermine is no good: as we've seen, Moorean responses to academic skeptical arguments, if they work at all, work only against unmotivated Academic skepticism. Any case against the Trust Requirement must not depend upon our ordinary normative judgments. Naïve replies, such as Cruelty 1-3, won't do.

However, to any anti-skeptic who reasonably rejects the Trust Requirement—and whose grounds for doing so are independent of his ordinary normative judgments, such as Cruelty-2—the skeptical arguments put forth against views of types (i) and (iii) are just unmotivated Academic skepticism. The question, then, is whether there are such grounds for rejecting the Trust Requirement. But it's hard for me to see any such grounds: the Trust Requirement is a core truth about reasoning. I doubt that anti-skeptics should take their stand here. That is not to say that normative skepticism is inevitable: it is just to say that it is the inevitable result of accepting Explanatory Independence.

4.7   WHO IS THE TARGET? (PART II)

It is now time to consider who, exactly, can reject Explanatory Independence. Explanatory Independence consists of three sub-theses:

(1) Our normative judgments do not explain why the normative facts are what they are.

(2) We do not have the particular normative faculties that we do because of the normative facts.

(3) There is no third factor that explains both why we have the normative faculties that we do and why the normative facts are what they are.

90

One way to deny Explanatory Independence is to deny (1) and embrace **constructivism**: the view that our normative judgments constitutively explain why the normative facts are what they are. Constructivism is a meta-normative view with a distinguished pedigree; I do not have anything novel to say about here, except to note that it can avoid the skeptical threat from Undermining 1-5. A second way out is to deny (2) and say that the normative facts causally explain why we have the normative faculties that we do. This strategy requires embracing two theses: first, **naturalism** or the view that normative facts are identical to natural facts (and thus causally efficacious);[55] second, that normative facts are well-positioned to explain why we have the normative faculties that we do.

So naturalists and constructivists can resist our skeptical argument, but only if they can endorse the claim that *unmotivated* Academic skeptical arguments fail – since, to foes of Explanatory Independence, Undermining 1-5 is an instance of unmotivated Academic skepticism. That is, skeptics will reply to attempts to block Undermining 1-5 by embracing constructivism or naturalism by pointing out that your judgment that constructivism or naturalism is true is itself a normative judgment. As a result, skeptics will insist, Undermining 1-5 shows that you shouldn't trust your judgment that constructivism or naturalism is true. Mooreans have no problem here, but philosophers who are less sanguine about Mooreanism have some explaining to do: how can we reject the skeptic's undermining hypothesis without somehow "begging the question" against her?[56] Since I'm happy with Moorean responses to skepticism, I'll set this worry aside; non-Mooreans, though, are in trouble regardless of their meta-normative view.[57]

Naturalism and constructivism offer two ways out; is there a third, namely to reject (3) and insist on a common factor explanation of normative facts, on the one hand, and why we have the normative faculties that we do, on the other? The only plausible candidates for the common

---

[55] I'll ignore the possibility that normative facts might be non-natural and yet causally efficacious.

[56] Will it help to say that the judgment that either constructivism or naturalism is true isn't a normative judgment but rather a meta-normative judgment, and thus isn't called into doubt by Undermining 1-5? No. Skeptics can run a version of Undermining 1-5 that focuses on our faculties for meta-normative judgments (using the premise that there is no explanatory connection between the meta-normative facts and our faculties for making meta-normative judgments).

[57] One horn of the "Darwinian Dilemma" presented in Street (2006) amounts to saying that naturalists must embrace a Moorean response to (unmotivated) Academic skepticism.

factor are the natural "realizers" of normative facts. Even those who deny that facts such *as setting a cat on fire is wrong* are causally efficacious can say first, that the act is wrong *because* it causes more pain than pleasure (say) and second, that we evolved a tendency to judge such actions wrong *because* they cause more pleasure than pain. This sort of common factor explanation is amenable even to non-naturalistic realists, who deny the causal efficacy of normative facts. [58] This strategy requires two claims: first, that the normative facts are true *because of* certain natural facts; second those same natural facts are well suited to explain why we have the normative faculties that we do. Can non-naturalistic realists make both those claims?

They cannot. A problem arises when we turn from particular normative claims—e.g. *setting a cat on fire is wrong*—to general normative principles, e.g. *causing more pain than pleasure is wrong*. The common factor strategy stumbles in dealing with general normative principles in two ways. First, natural facts can explain the wrongness of a particular act by showing that it is an instance of a more general kind of act (e.g. an act of causing more pain than pleasure)—and thereby showing that a general rule applies to it—but *that* sort of move cannot explain why the general rules are what they are (e.g. it cannot explain *why* causing more pain than pleasure is wrong). But it's hard to see any other explanatory strategies available to realists for using natural facts to explain normative facts; if there aren't any, there is no way to use natural facts to explain why the general rules are what they are. Second, general normative rules are necessary truths, so if there are natural truths that explain them, then there are natural truths that explain them in all worlds. But what truths are these supposed to be? I'm at a loss; I just can't see any necessary natural facts that explain why causing more pain than pleasure is wrong. So the third approach doesn't work: it leads to skepticism about general normative rules.

One response to this difficulty is to deny that there are general normative rules, or that they play any important role in our normative reasoning. This is a version of **particularism**, the view that there is no reason to suppose that there are true moral principles, where moral principles must be codifiable: not infinitely complex or applicable only to single cases. [59] The

---

[58] Enoch (forthcoming) defends a version of non-naturalistic realism on roughly these grounds.

[59] McKeever and Ridge (2006) distinguish between four versions of particularism; they call the version that I discuss here "Principle Skepticism Particularism" (15). Dancy (2006) defends a stronger version of particularism: he denies that there are moral principles (as opposed to denying that we ought to expect, in advance of doing normative ethics, to find any).

good news is that particularists can shrug off skepticism about normative generalities. The bad news is that particularism entails that neither the normative facts nor their naturalistic realizers can explain why we have the normative faculties that we do. We can explain why creatures inclined to judge actions conducive to total happiness, or the preservation of human life, or whatever would be more likely to survive; we could not do so for an inclination to believe in a shapeless or non-codifiable set of facts.[60] Returning to the example of cat immolation: it is plausible that we evolved the normative faculties that we did because judging that actions that cause more pain than pleasure are wrong has some survival value; if the fact that the action causes more pleasure than pain just is the wrongness of setting the cat on fire then the normative facts explain our normative judgments. It is not plausible, however, that we evolved the normative faculties that we did because judging that actions with the highly specific and local set of features borne by this instance of cat-immolation are wrong: it is vanishingly unlikely that any of my ancestors ever encountered an action with those features. None of this is to say that there must be a single true moral principle, or that the principles need to be as simple as *actions that cause more pain than pleasure are always wrong*; the point is just that the principles must be general enough that a tendency to judge in accordance with them could have survival value.[61]

In sum, then, non-naturalistic realists face a dilemma: do general normative principles play an important role in our reasoning? If yes, they are in trouble because there is no explanatory connection between normative generalities and our attitudes. If no, they are in trouble because non-codifiable normative principles are ill-suited to explain why we have the attitudes that we do. Neither answer is consistent with rejecting Explanatory Independence by rejecting (3).

---

[60] One thesis that sometimes goes by the name "particularism" is that we should only expect to find moral principles involving "thick" normative terms such as *cruelty* or *cowardice* – we should not expect to find moral principles statable using only normatively neutral language. I do not mean for what I have said here to be an attack on this thesis. One difficulty in arguing against this weaker version of particularism is that "thick" moral terms tend to be *psychological* terms, and determining the relationship between the psychological and the natural is—to put it mildly—beyond the scope of this paper.

[61] I've stated particularism as a thesis about *moral* principles. Fans of particularism about epistemology would also be unable to reject Explanatory Independence, but as far as I know epistemic particularism has no adherents (it's interesting to ask why not, though I confess that I am at a loss to explain the asymmetry).

4.8   CONCLUSION: AUSTERITY WITHOUT EMPIRICISM

In this chapter I have reached three conclusions. First, I have argued that the best version of the worry about the causal origins of our normative judgments—Undermining 1-5—is a rare and powerful type of skeptical argument. It is more effective than Coincidence 1-5 because it does not just raise a puzzle about our normative judgments but rather undermines their support. It is more effective than standard Academic skeptical arguments because it gives motivation for its undermining hypothesis – and, crucially, fans of Explanatory Independence accept the premises of this argument. Undermining 1-5 is thus a rare and devastating breed: a motivated Academic skeptical argument.

Second, I have argued that adopting an alternative epistemology of the normative—a view of type (i)-(iii) wherein the output of our faculties is not evidence for normative conclusions—cannot defeat skepticism. A skeptic armed with the Trust Requirement can make trouble for alternative approaches to normative epistemology.

Third, I have suggested a way out: insist on explanatory relations between normative facts and our faculties of normative judgment. We can do this by adopting constructivism and saying the normative facts are what they are because we are inclined to make certain judgments, or by adopting naturalism and identifying normative facts with some of the natural facts that explain how we came to have our normative faculties. Against these positions, Undermining 1-5 is a form of *unmotivated* Academic skepticism. If we do insist on explanatory relations between normative facts and our normative faculties, we must reject both particularism and non-naturalistic realism.

If we step back far enough, the argument of this paper resembles a familiar sort philosophical worry: how can we know about "invisible" facts - about norms, or modal properties, or abstract objects? We can't, after all, learn about them by perceiving them, or by talking to people who perceived them, or building instruments to measure them. Philosophers have tried to make this sort of worry precise in many ways: through the empiricist account of concept formation (i.e. the "copy principle"), verificationism, the causal theory of knowledge, and Harman's explanatory relevance test. These attempts are all versions of empiricism, giving pride of place to *experience*. This means that whether or not we can defeat these arguments on

their own terms—accusing them of having outlandish implications, say, or of being self-defeating—there is always room to resist them by just denying empiricism. This can shut down the familiar arguments against "invisible" facts: though empiricism's emphasis on experience is pretty plausible, it's difficult to really argue for the copy principle, or verificationism, or the causal theory of knowledge, or the explanatory relevance test. Hume, Ayer, Benacerraf, and Harman—to cite perhaps the most prominent proponents of the aforementioned strategies—just announce these principles; they don't argue for them.[62] While my argument against a kind of "invisible" fact—normative facts, understood along non-naturalistic realist lines—shares an aim with these earlier attempts, it doesn't share their empiricism. Nowhere do I insist on experience as a royal road to justification. Rejecting empiricism leaves my argument untouched.

That doesn't mean I make no assumptions – just that I make different ones. My argument doesn't give pride of place to experience, but rather gives it to inference to the best explanation. The thought that inference to the best explanation is always rational is the key premise in my argument for the Explanatory Constraint. While jettisoning inference to the best explanation entirely is surely too radical, perhaps one can restrict its applicability; I suspect that this is the best line of resistance to the argument of this paper. The alternative picture rejects the application of the Explanatory Constraint to the normative domain (which, in turn, requires rejecting that inference to the best explanation is always rational in the normative domain). On this alternative picture there is no reason to conclude, from the lack of explanatory connection between the normative facts and the output of our faculties, that the evidential force of the output of our faculties is undermined. This alternative would still take some philosophical work: for one thing, we would need an alternative picture of evidential support wherein the output of normative faculties *is* evidence for normative claims. All told, I do not accept this alternative: I'd need a compelling positive story about how it is that our normative judgments are evidence for normative claims. But this alternative position is tempting nonetheless, so there is another way out.

---

[62] See Hume (2000), Ayer (1952), Benacerraf (1973), and Harman (1989).

BIBLIOGRAPHY

Ayer, A (1952). *Language, Truth, and Logic*. New York: Dover.

Benacerraf, P (1973). Mathematical Truth. *The Journal of Philosophy* 70, 661-679.

Christensen, D (2007). Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals. *Oxford Studies in Epistemology*, 2, 3 – 31.

Cohen, S (2002). Basic Knowledge and the Problem of Easy Knowledge. *Philosophy and Phenomenological Research*, 65(2), 309-329

Dancy, J (2006). *Ethics Without Principles*. Oxford: Oxford University Press.

Dworkin, R (1996). Objectivity and Truth: You'd Better Believe It. *Philosophy and Public Affairs* 25: 87-139.

Enoch, D (Forthcoming). The Epistemological Challenge to Metanormative Realism. *Philosophical Studies*.

Evans, G (2000). Reference and Contingency. In *Collected Papers*. New York: Oxford University Press.

Field, H (1991) *Realism, Mathematics, and Modality*. Oxford: Blackwell.

Harman, G (1989) Ethics and Observation. In Sayre-McCord (1989)

Hawthorne, J (2002). Deeply Contingent Apriori Knowledge. *Philosophy and Phenomenological Research,* 65(2), 247-269.

Hawthorne, J (2005). The Case for Closure. In Steup, M and Sosa, E (Eds.) *Contemporary Debates in Epistemology*. Oxford: Blackwell.

Hume, D (2000). *A Treatise of Human Nature*. Oxford: Oxford University Press.

Joyce, R (2006). *The Evolution of Morality*. Cambridge: MIT Press.

Kotzen, M (ms). A Formal Account of Evidential Defeat.

McKeever, S and Ridge, M (2006). *Principled Ethics*. Oxford: Oxford University Press.

Nagel, T (1986). *The View from Nowhere*. Oxford: Oxford University Press.

Nozick, R (1981) *Philosophical Explanations*. Cambridge: Harvard University Press.

Parfit, D (ms). *Rediscovering Reasons*.

Pryor, J (2000). The Skeptic and the Dogmatist. *Nous*, 34, 517-549.

Pryor, J (2004). What's Wrong with Moore's Argument? *Philosophical Issues*, 14, 349-78.

Pryor, J (2008). When Warrant Transmits. In Coliva, A Coliva, Annalisa (Ed.) *Mind, Meaning and Knowledge: Themes from the Philosophy of Crispin Wright*. Oxford: Oxford University Press.

Pryor, J (ms). Problems for Credulism.

Quinn, W (1986). Truth and Explanation in Ethics. *Ethics* 96: 524-544.

Raz, J (1999). *Practical Reason and Norms*. Oxford: Oxford University Press.

Sayre-McCord, J (1989). *Essays on Moral Realism*. New York: Cornell University Press.

Scanlon, T (2000). *What We Owe to Each Other*. Cambridge: Belknap Press.

Schafer, K (Forthcoming). "Evolution and Normative Skepticism," *Australasian Journal of Philosophy*.

Schiffer, S (2004). Skepticism and the Vagaries of Justified Belief. *Philosophical   Studies* 119, 161-184.

Setiya, K (Forthcoming). "Does Moral Theory Corrupt Youth?" *Philosophical Topics*.

Silins, N (2008). Basic Justification and the Moorean Response to the Skeptic. *Oxford Studies in Epistemology*, 2, 108-142.

Street, S (2006). A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies* 127: 109-166.

Street, S (ms). Objectivity and Truth: You'd Better Rethink It.

Street, S (forthcoming). Evolution and the Normativity of Epistemic Reasons. *Oxford Studies in Metaethics*.

Sturgeon, N (1989). Moral Explanations. In Sayre-McCord (1989).

Weatherson, B (2007). The Bayesian and the Dogmatist. *Proceedings of the Aristotelian Society*, 107, 169-85.

Weisberg, J (Forthcoming). Bootstrapping in General. *Philosophy and Phenomenological Research*.

White, R (2006). Problems for Dogmatism. *Philosophical Studies*, 131, 525-557.

Wright, C (1985). Facts and certainty. *Proceedings of the British Academy*, 71, 429–72.

Wright, C (2002). (Anti-) Skeptics Simple and Subtle: Moore and McDowell. *Philosophy and Phenomenological Research*, 65, 330-48.

Wright, C (2004).   Warrant for Nothing (and Foundations for Free)? *Aristotelian Society Supplement*, 78, 167-212.

Wright, C (2008). The Perils of Dogmatism. In Nuccetelli, S & Seay, G (Eds.) *Themes from G. E. Moore: New Essays in Epistemology and Ethics*. Oxford: Oxford University Press.